AD__________________
(Leave blank)

Award Number: **W81XWH-04-1-0142**

TITLE: VITAL (Vanguard Investigations of Therapeutic Approaches to Lung Cancer)

PRINCIPAL INVESTIGATOR:
Waun Ki Hong, M.D., Principle Investigator
Reuben Lotan, Ph.D., Co-Principle Investigator
David Stewart, M.D., Co-Principle Investigator

CONTRACTING ORGANIZATION:
The University of Texas M.D. Anderson Cancer Center
Houston, TX 77030

REPORT DATE: January 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

☐ Distribution limited to U.S. Government agencies only; report contains proprietary information

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 14-01-2009	**2. REPORT TYPE** Annual	**3. DATES COVERED** *(From - To)* 15 Dec 2007 – 14 Dec 2008

4. TITLE AND SUBTITLE VITAL (Vanguard Investigations of Therapeutic Approaches to Lung Cancer)	**5a. CONTRACT NUMBER** W81XWH-04-1-0142
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Waun Ki Hong, M.D. Reuben Lotan, Ph.D. David Stewart, M.D.	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Texas M.D. Anderson Cancer Center Houston, TX 77030	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The VITAL Research Program will provide a better understanding of the cellular and molecular processes that drive lung tumorigenesis so that an accurate risk model for recurrence and/or the development of the secondary primary tumor can be developed, and the biologic agents most effective in reducing these events in the group of high-risk patients can be identified. We will be incorporating retrospective clinical trial specimens to develop our risk model and validating it with specimens collected from our Vanguard study. The research projects are proceeding well as proposed, producing valuable findings with cell lines, and will validate these results using the clinical samples obtained from the VITAL trials in the coming years.

15. SUBJECT TERMS

Lung cancer, risk model, cancer recurrence, clinical trials

16. SECURITY CLASSIFICATION OF: U			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	**309**	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

INTRODUCTION 2
BODY 3

- Project 1 3
- Project 2 6
- Project 3 21
- Project 4 35
- Project 5 44
- Core B Biostatistics and Data Management 50
- Core C Pathology and Specimen Procurement 52

KEY RESEARCH ACCOMPLISHMENTS 58

REPORTABLE OUTCOMES 60

CONCLUSIONS 64

REFERENCES 65

APPENDICES 67

Appendix 1: (Core B) Trial Summary via Event Charts and Data Management Activities for VITAL

Appendix 2: Publications

INTRODUCTION

Smoking-related cancers such as lung and head and neck cancers are a major cause of cancer death in the United States. About 25% of lung cancer patients are diagnosed with stage I or II disease and undergo surgery with curative intent, but the 5-year survival for this group of patients is only 30%-70%. Patients with a strong history of smoking and prior early-stage cancer are found to be at high risk for cancer recurrence or development of second primary tumors (SPTs). An effective adjuvant therapy after surgery in this group of patients is not well established yet. The survival benefit of adjuvant chemotherapy was uncertain until recent findings reported by Winton and colleagues (Winton et al., 2005). They found that adjuvant chemotherapy (vinorelbine and cisplatin) increases the 5-year survival of surgically resected non-small cell lung cancer (NSCLC) patients, resolving the debate over the benefit of adjuvant chemotherapy. Thus, better-designed clinical trials and basic research are needed to establish the standard of care for these patients after surgery.

The program VITAL (Vanguard Trial of Investigational Therapeutics in Adjuvant Treatment of Lung Cancer) initiated in 2003 was developed to gain a better understanding of the molecular events underlying the progression of NSCLC in order to develop a risk model for cancer recurrence and development of smoking-related SPT in the high-risk population, and to identify effective preventive agents for this group of patients. Specifically, our objectives are:

- To identify biologically-based treatments for prevention of cancer recurrence and development of second primary tumors in high-risk patients;
- To understand molecular events in premalignant tissues that contribute to progression or malignancy;
- To develop a risk prediction model for disease recurrence and development of second primary tumors in high-risk patients by combining clinical treatment outcomes with molecular and imaging data.

Three clinical trials were proposed, in part, to acquire the necessary correlative samples to develop this risk model, which will significantly improve decision-making for patients and physicians in the management of this challenging disease. Histologic assessment was planned to determine whether malignant changes would occur during this time period. Despite substantial efforts, our patient accrual was significantly lower than expected due to a number of factors; thus, a ReVITALization plan was proposed (see revised Aims below) and approved by the DoD in the previous funding period. Implementation of the alternative ReVITALization strategy over the past year was based on the revised project aims that were developed to accomplish our goal of the development of the risk model. An overview of the changes is provided below with additional details in each relevant project. It should be noted that work described in Project 1 related to these reviewed aims is planned to continue until January 2010; a request for an unfunded 12-month extension has been submitted for approval to the DoD.

ReVITALization Aims:

1. **Circumvent low accruals using surgical specimens in our tissue bank.** These specimens (about 500 samples) of resected lung cancer will be utilized for biomarker assessment and will serve as the foundation for a biomarker-based risk assessment model.

2. **Continue enrolling patients in our Vanguard trial to accrue 50-60 patients.** This cohort will provide sufficient biospecimens for the aims proposed in the other projects of the VITAL program. Additionally, the clinical data obtained from these patients will be

used to test the biomarker-based risk assessment model and the follow-up bronchoscopy specimens will provide important information for biomarker changes in the bronchial epithelium.

3. **Close the celecoxib and erlotinib trials to focus resources on specimen analyses to develop the biomarker risk model.**

4. **Perform two additional discovery projects related to increased risk that are only now possible due to continued progress in VITAL.**
   - **a.** Identify gene expression signatures in bronchial brush specimens using high-throughput genomics approach.
   - **b.** Identify genes expression signatures in epithelial cells detected by LIFE bronchoscopy that determine aggressiveness.

This report summarizes work conducted over the past year of the research period, highlights key research accomplishments and reportable outcomes with the bibliography of all publications and meeting abstracts derived from VITAL during this timeframe.

PROGRESS REPORT (BODY)

Project 1: Biologic Approaches for Adjuvant Treatment of Aerodigestive Tract Cancer

(PI and co-PIs: Drs. Waun Ki Hong, Edward S. Kim, Rodolfo C. Morice, David J. Stewart)

Aim 1 Assess the smoking-related disease-free survival in patients who are current or former smokers with a prior definitively-treated stage I/II lung or head and neck cancer.

The main objective for this project was to open the Vanguard study at MDACC as well as the 2 other participating sites. Enrollment was planned for a total of 300 patients with definitively treated stage I/II lung or head and neck cancer and at least a 20-pack-year smoking history. Patients undergo baseline testing including chest x-ray, CT scan, labs, bronchoscopy, and other specimen collections (i.e., sputum, saliva, serologies). Bronchoscopies and specimen collection are performed at baseline and at months 12, 24, and 36. White-light alone or white-light and autofluorescence modalities are used. Abnormal areas detected by bronchoscopy are biopsied. Histologic assessment is performed to determine whether malignant changes will occur during the time period. If severe dysplasia, carcinoma *in situ,* or carcinoma is discovered, patients follow the plans outlined in the clinical protocol. Once patients have completed 3 years of testing, they are followed until the study is completed. As per the revised ReVITALization aims, the study will be closed when a total of 50-60 patients have been accrued; all patients will be followed as outlined above.

Summary of Research Findings

In the past year, we have continued enrolling patients in the VITAL/Vanguard trials. We have accomplished our goal for enrollment (at least 50 patients) and plan to close the study to new patients after January 2009. A total of 53 patients have been enrolled in the Vanguard trial, with two additional patients to be enrolled by the end of February 2009 (55 total patients). Currently, 31 patients have completed both the baseline and 12-month bronchoscopy. Nine additional bronchoscopy procedures are scheduled to be completed by January 2010, which will bring our number of evaluable patients to at least 40 by the end of the next year. Patient clinical data and tissues have been and continue to be collected and will be distributed to investigators of VITAL research projects through the VITAL Pathology Core. Over the next 12 months, the Pathology Core will complete the immunohistochemistry (IHC) analysis of the tissue microarray (TMA) set for Project 1. For details, please refer to the Pathology Core update (pg. 49). As detailed in the report from the Biostatistical Core (see pg. 47), this sample size will allow us to update the ReVITALization database with patient data for analysis to support this collective research effort. New statistical methods are being used to evaluate the interactions for combination therapy to determine if the effects are synergistic, additive, or antagonistic.

Plan: Close the Vanguard trial to accrual.

Following completion of accrual in January 2009 and subsequent patient follow-up, we will have the necessary prospective specimens to fulfill the goals of Projects 2-5. As noted in the revised aims, the planned biomarker analyses in these projects will be supplemented using retrospective specimens to maximize data acquisition while minimizing the time required to develop the proposed risk model (See Aim 3 and the Core C report below).

Aim 2 Evaluate effects of biologic agents as adjuvant therapy on the modulation of histology and specific biomarkers in this high-risk population.

Current adjuvant chemotherapy offers some benefits in the high-risk patients, but is not a long-term preventive strategy. Our plan was to open several biologic adjuvant clinical trials with novel agents such as celecoxib, erlotinib, lonafarnib, and possibly others; however, poor accrual due to changes in the standard of care for lung cancer patients prohibited successful completion of the proposed trials and, thus, the trials were abandoned as previously described. Our revised aim is now focused on the timely development of the risk model (see Aim 3 / Revised Task 3). The retrospective specimens will be used for the proposed analyses and correlated with the available clinical data.

Summary of Research Findings

As noted previously, the celecoxib trial was closed, and the proposed erlotinib trial was deferred. We will continue to focus our efforts on the productive analysis of acquired samples from the clinical trial and from our tissue bank, leading to the timely development of a lung cancer risk model.

Aim 3 Develop a lung cancer risk model to help predict the likelihood of development of relapse or new smoking related primary tumors

Patients with a history of smoking and a prior surgically resected stage I/II head and neck or lung cancer are at high risk for cancer recurrence or SPTs. There are no standard interventions that have been proven to help reduce the risk of cancer occurrence. A Gail risk model implemented in the initial management of breast cancer screening has proven useful and has helped with early detection and more stringent follow-up in the higher risk cohorts. Patients enrolled in the Vanguard trial will have aggressive post-operative follow-up with analysis including frequent serologies, bronchial specimens and CT scanning. Trends in these multiple biomarkers will be analyzed and used to develop a predictive model. Establishing a risk model will eventually help identify patients who may be at higher risk for lung cancer development and promote earlier interventions for prevention.

As noted previously, we have revised Aim 3 to develop a lung cancer risk model to help predict the likelihood of cancer recurrence and second primary tumor (SPT) development utilizing clinical, pathologic and biomarker information obtained prospectively and retrospectively from the high-risk population of patients (Years 4-5).

Summary of Research Findings

Dr. Ignacio Wistuba, Pathology Core Director, has identified archival tissue specimens in our tissue bank from over 600 surgically resected lung cancers, stages I/II, acquired over the period of 2002-2005. These specimens are available for our use and meet the criteria in our proposed biomarker studies. These retrospective specimens all have follow-up data for a minimum of 2 years. All NSCLC case specimens with adjacent bronchial structures will be included in the study. Complete clinical and pathologic characteristics will need to be audited and incorporated into our shared VITAL/PROSPECT database**.** We will thus increase our samples size from 300 to 500 for use in developing our risk model by supplementing the available prospective specimens with additional retrospective specimens*.* Our focus in VITAL is on lung epithelial changes, rather than changes in the tumor itself; hence, by including all patients with specimens from adjacent bronchial epithelium, the goal of developing a risk model can be achieved. (See Core C for more detail.)

We have completed entry into the pathology database for the 500 archival tissue specimens of resected lung cancers. The clinical database has been developed and data entry is in the process of being completed via both patient record review and contacting the patient or family members directly. As permission must be granted first before person-to-person contact can be made, more than 300 IRB-approved letters to patients have been mailed requesting permission to contact them regarding this study. Once we receive patient permission, we will contact individual patients directly for a long-term update on their condition. Patients who continue to be followed at MDACC and have adequate records will not need to be contacted through this mechanism.

Key Research Accomplishments

- Enrolled 9 more patients in the Vanguard study and have 2 patients pending, for a total of 55.
- Continued to collect patient clinical data and tissues for distribution to support research projects in the VITAL grant.
- 31 patients have completed the 12-month mark/bronchoscopy, with a projected total of at least 40 evaluable patients by the end of January 2009.
- ReVITALization plan is being implemented with data entry into the pathology database related to the 500 archived tissue specimens, and updating of the clinical database for these patients.

Meeting Abstracts

We plan to report clinical findings on the Vanguard Trial at a major meeting in 2010.

Conclusion

The completion of the Vanguard trial is important to this grant. However, the ultimate goal is the development of a risk model for development of SPT and recurrence. This model will be developed utilizing the resources from both the Vanguard trial and our Lung Cancer Tissue Bank as a top priority for our program.

Project 2: Identification of Biomarkers of Response to Chemoprevention Agents in Lung Epithelium

(PI and co-PIs: Li Mao, M.D., Reuben Lotan, Ph.D., John Minna, M.D.)

Lung cancer continues to be the deadliest among all cancers in the United States with over 165,000 deaths annually for the last few years and an overall 5-year survival rate of less than 15% (Jemal et al., 2007). Early detection of premalignant lesions or tumors appears to be an efficient approach to reducing the morbidity and mortality from lung cancer because the survival of early stage lung cancer patients is much better than that of patients with advanced cancers. Therefore, new strategies for the early diagnosis, prevention and treatment of this dreadful disease are urgently needed (Wistuba and Gazdar 2006; Sato et al., 2007). The development of early detection tools for lung cancer requires improved molecular testing by identification and understanding of early events in the multi-step process of lung carcinogenesis, which involves the accumulation of genetic and epigenetic alterations over the long course of exposure to carcinogens such as tobacco smoke (Mao, 2002; Wistuba and Gazdar 2006; Sato et al., 2007). To date, there are no validated biomarkers for early detection. Moreover, one or a few genes may not provide sufficient specificity given the multi-factorial process of lung carcinogenesis and heterogeneous nature of lung cancer. Thus, the effort to search for more specific and sensitive biomarkers of early lung cancer is warranted. The development of high-throughput gene expression analyses, e.g., DNA-chips or microarrays, provides opportunities to define biomarkers (signatures) of risk of cancer development. During the last few years, several studies reported molecular classification of human lung carcinomas on the basis of gene expression and described numerous putative biological markers of cancer (Meyerson et al., 2004). However, only limited number of studies has attempted to identify genes that are modulated at early stages of human lung carcinogenesis such as premalignant state because of the limited availability of premalignant lung tissues suitable for RNA extraction. We hypothesized that immortalized, transformed and tumorigenic human bronchial epithelial cell (HBEC) line models will have similar abnormalities in gene expression profiles as premalignant and malignant tissues *in vivo*. Therefore, such cell models will be useful to identify markers of early disease.

We proposed to use genomic and proteomic analyses to identify changes in gene expression (including mRNA and miRNAs) and proteins which correlate/associate with cancer risk in the carcinogen damaged aerodigestive tract field and also use these signatures to monitor the response of this field to chemoprevention. We will develop and use a model HBEC system to study the effect of specific oncogenic changes and also the response of these manipulated HBECs to various carcinogenic and chemoprevention agents. Thus, we will determine modifications of these changes by chemopreventive agents in premalignant cells *in vitro* and to use probes for the modified genes and proteins to analyze tissue specimens from individuals participating in the chemoprevention clinical trials.

Aim 1 Develop immortalized human bronchial epithelial cell cultures using a subset of patient tissue specimens collected in Project 1 and characterize the expression profiles of these cells using oligonucleotide based microarrays.

The main goal of this aim of this project is to establish these cultures from lung cancer patients and persons without lung cancer, including those patients entered onto the clinical trial described in Project 1, and to characterize their gene expression profiles.

Summary of Research Findings

A. Generating immortalized HBECs and Small Airway Epithelial Cells (HSAECs) from different individuals that were collected on various tissue procurement protocols at The University of Texas Southwestern and at M. D. Anderson Cancer Center.

To date, we have generated HBECS from over 45 different individuals and have recently immortalized 10 peripheral small airway epithelial cells (HSAECs) (Fig. 1).

These specimens represent males, females, a spectrum of smoking status, as well as persons with and without lung cancer. We have introduced oncogenic exposures into several of the HBECS and HSAECs, such as oncogenic KRASV12 and EGFR with mutations, that are found in patients. We have also immortalized 15 of these same bronchial epithelial specimens with oncogenic HPV E6 and E7, and made several pairs of immortalized HBECs and lung cancer cell lines from the same patient.

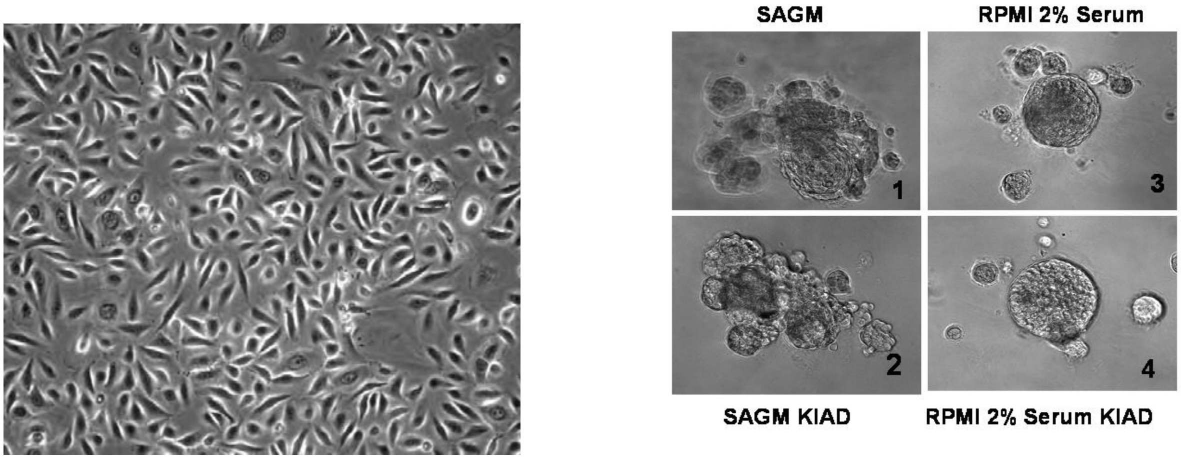


Figure 1. Immortalized normal human small airway epithelial cells (HSAEC) growing in matrigel under various growth conditions. Human airway epithelial cells from the peripheral lung were immortalized with hTERT and CDK4 and grown in defined small airway growth media (SAGM) in monolayer culture (top panel). They were then plated in matrigel and tested for growth and differentiation in SAGM or in RPMI1640 medium with 2% fetal calf serum with or without a supplement to induce differentiation (KIAD: K, KGF/FGF-7; I, 3-isobutyl-1-methylxanthine (IBMX); A, 8-bomoadenosine 3',5'-cyclic monophosphate; D, dexamethasone). In matrigel, the HSAECs undergo branching morphogenesis and KIAD aids to the differentiation, which is partially reversed by serum.

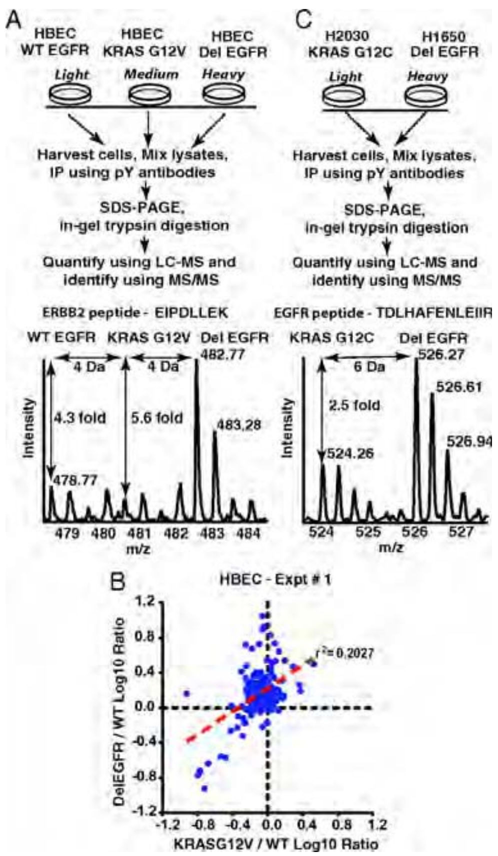


Figure 2. Increased tyrosine phosphorylation of proteins in HBECs or adenocarcinoma cells expressing mutant EGFR compared with cells expressing mutant KRAS. Increased tyrosine phosphorylation of proteins in HBECs or adenocarcinoma cells expressing mutant EGFR compared with cells expressing mutant KRAS. Schematic of experimental design and representative MS spectra of a peptide of ERBB2 (A) and a peptide of EGFR (C) identified in the phosphotyrosine immunoprecipitates of proteins from lysates of HBECs and adenocarcinoma cell lines, respectively. (B) The degree of tyrosine phosphorylation of proteins in mutant RAS expressing cells does not correlate with that in mutant EGFR-expressing cells as demonstrated in the log-log plot of the two ratios obtained in this experiment (Del EGFR/WT EGFR and Mut RAS/WT EGFR). Log 10 transformation of SILAC ratios for individual proteins and linear regression analysis yields a coefficient of determination (r2) of 0.20, and a Pearson product-moment correlation (r) of 0.45. (Guha U., et al., PNAS 2008;105:14112-14117).

B. Developed new ways and media to have the airway epithelial cells differentiate and study their stem cell like properties. We have used the defined small airway growth medium (SAGM) combined with Matrigel culture to study the ability of HSAECs to differentiate *in vitro*. In matrigel, they form cyst-like structures and show increased expression of CC10 and SP-C. This approach has also involved adding differentiation factors (KIAD, see legend of Fig. 1) to the SAGM or to serum containing media. The addition of serum also promotes differentiation with increased expression of Type 2 cell marker SP-C and Type 1 cell marker Aquaporin 5. The HSAECs do not express Type 2 cell markers SP-A, SP-C by qRT- PCR, but do express low levels of Clara cell marker CC10, and high levels of Notch stem cell markers. The HSAECs have a high percentage of cells that are aldefluor-positive as well as express high levels of Notch stem cell genes (both markers of stem cells).

C. Induced oncogenic changes in the HSAECs and performed genomic analyses. We have introduced various oncogenic changes into the HSAECs beginning with the introduction of oncogenic KRASV12, which causes the cells to lose some of their differentiating ability. These “preneoplastic” HSAECs are being tested with chemoprevention agents as well as having microarray mRNA expression profiles and array-based CGH copy number profiles performed.

D. Identifying a large number of new tyrosine phosphorylation targets of oncogenes including KRASV12 and mutant EGFR in HBECs using unbiased phosphoproteomic approaches combined with quantitative analyses with mass spectroscopy. In collaboration with the laboratory of Dr. Harold Varmus at Memorial Sloan Kettering Cancer Center (MSKCC), we studied the tyrosine phosphorylation changes induced in HBECs by the presence of mutant oncogenes. We have used unbiased phosphoproteomic approaches, based on quantitative mass spectrometry using stable isotope labeling with amino acids in cell culture (SILAC), to identify tyrosine phosphorylated proteins in isogenic human bronchial epithelial cells (HBECs) and human lung adenocarcinoma cell lines, expressing either of the two mutant alleles of *EGFR* (*L858R* and *Del E746-A750*), or a mutant *KRAS* allele, which are common in human lung adenocarcinomas (Fig. 2). The mass spec results were validated by standard Western blotting experiments (Fig. 3).

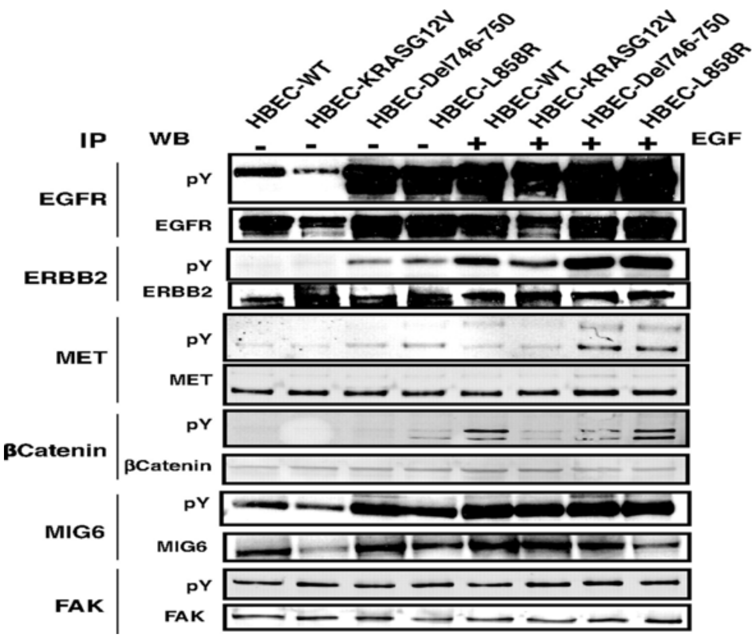


Figure 3. Validation of mass spectrometry-based quantitation by immunoprecipitation and Western blots of representative proteins. Validation of mass spectrometry-based quantitation by immunoprecipitation and Western blots of representative proteins. Immunoprecipitation was done with antibodies to the indicated proteins from lysates of HBECs, and Western blots done with anti-phosphotyrosine and protein-specific antibodies. (Guha U., et al., PNAS 2008;105:14112-14117).

Tyrosine phosphorylation of signaling molecules was greater in HBECs expressing the mutant EGFRs than in cells expressing wild type (WT) EGFR or mutant KRAS. Receptor tyrosine kinases (such as EGFR, ERBB2, MET, and IGF1R), and Mig-6, an inhibitor of EGFR signaling, were more phosphorylated in HBECs expressing mutant EGFR than in cells expressing WT EGFR or mutant RAS. Phosphorylation of some proteins differed in the two EGFR mutant-expressing cells; for example, some cell-junction proteins (β-catenin, plakoglobin, and E-cadherin) were more phosphorylated in HBECs expressing L858R EGFR than in cells expressing Del EGFR. There were also differences in degree of phosphorylation at individual tyrosine sites within a protein; for example, a previously uncharacterized phosphorylation site in the nucleotide-binding loop of the kinase domains of EGFR (Y727), ERBB2 (Y735), or ERBB4 (Y733), is significantly more phosphorylated in HBECs expressing the deletion mutant than in cells expressing the wild type or L858R EGFR. Signaling molecules not previously implicated in ERBB signaling, such as polymerase transcript release factor (PTRF), were also phosphorylated in cells expressing mutant EGFR. Bayesian network analysis of these and other datasets revealed that PTRF might be a potentially important component of the ERBB signaling network (Fig. 4). This proteomic information provides important new biomarkers and potential targets for chemoprevention of these specific oncogenic changes in human airway epithelial cells, and provides a new model to identify similar markers after other oncogenic changes.

Aim 2 Characterize effects of the chemo preventive agents used in Project 1 on cell proliferation and apoptosis in the immortalized human bronchial epithelial cell cultures developed in Specific Aim 1.

We will determine the potential role of different chemo preventive agents [e.g., celecoxib, N-[4-hydroxyphenyl]retinamide (4-HPR), Iressa (gefitinib), and SCH63663] alone or in combination with one another for their effects on cell proliferation and apoptosis in cell cultures established in Aim 1. We will also determine the relative sensitivity among the various cell cultures to each of the agents by determining the 50% growth inhibitory concentration (IC_{50}).

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 3 Identify gene expression and protein “signatures” which reflect lung tumorigenesis and sensitivity or resistance to chemo preventive regimens proposed in Project 1, and to validate the signatures and to determine their biological importance in precancer cell models of lung cancer.

Summary of Research Findings

A. Perform high-density array CGH analysis of lung cancers to identify regions commonly amplified in lung cancer. In collaboration with the laboratories of Wan Lam at the Vancouver British Columbia Cancer Agency (BCCA) and Jon Pollack at Stanford University Medical Center, we have been studying DNA copy number changes in lung cancers and preneoplastic tissues, identifying the key genes within these amplicons and then performing functional studies to know which are the most important in the pathogenesis of lung cancer.

Chromosomal translocation is the best-characterized genetic mechanism for oncogene activation. However, there are documented examples of activation by alternate mechanisms, for example gene dosage increase, though its prevalence is unclear. Here, we answered the fundamental question of the contribution of DNA amplification as a molecular mechanism driving oncogenesis. Comparing 104 cancer lines representing diverse tissue origins identified genes residing in amplification 'hotspots,' we discovered an unexpected frequency of genes activated by this mechanism (Fig. 5). The amplification regions were correlated with microarray data and the overexpressed genes in each of the regions were identified (Fig. 6). The 3,431 amplicons identified represent approximately 10 per hematological and approximately 36 per epithelial cancer genome. Many recurrently amplified oncogenes were previously known to be activated only by disease-specific translocations. The 135 hotspots identified contain 538 unique genes and are enriched for proliferation, apoptosis and linage-dependency genes, reflecting functions advantageous to tumor growth. Integrating gene dosage with expression data validated the downstream impact of the novel amplification events in both cell lines and clinical samples. For example, multiple downstream components of the EGFR family signaling pathway, including CDK5, AKT1 and SHC1, are overexpressed as a direct result of gene amplification in lung cancer (Fig. 7). Our findings suggest that amplification is far more common a mechanism of oncogene activation than previously believed and that specific regions of the genome are hotspots of amplification. These amplicons provide important new biomarkers that may be present in preneoplastic lesions (see example below in Fig. 10) and could be promising targets for the monitoring and chemoprevention of lung cancer.

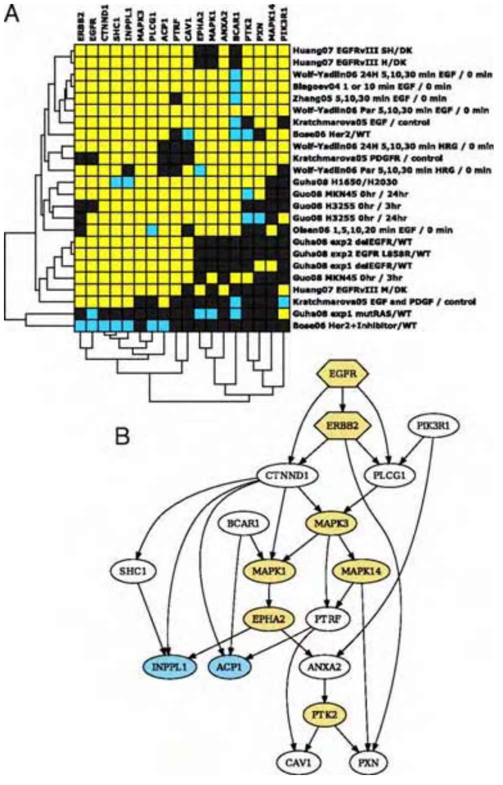

A
ERBB2
EGFR
CTNND1
SHC1
INPPL1
MAPK3
PLCG1
ACP1
PTRF
CAV1
EPHA2
MAPK1
ANXA2
BCAR1
PTK2
PXN
MAPK14
PIK3R1
Huang07 EGFRvIII SH/DK
Huang07 EGFRvIII H/DK
Wolf-Yadlin06 24H 5,10,30 min EGF / 0 min
Blagoev04 1 or 10 min EGF / 0 min
Zhang05 5,10,30 min EGF / 0 min
Wolf-Yadlin06 Par 5,10,30 min EGF / 0 min
Kratchmarova05 EGF / control
Bose06 Her2/WT
Wolf-Yadlin06 24H 5,10,30 min HRG / 0 min
Kratchmarova05 PDGFR / control
Wolf-Yadlin06 Par 5,10,30 min HRG / 0 min
Guha08 H1650/H2030
Guo08 MKN45 0hr / 24hr
Guo08 H3255 0hr / 3hr
Guo08 H3255 0hr / 24hr
Olsen06 1,5,10,20 min EGF / 0 min
Guha08 exp2 delEGFR/WT
Guha08 exp2 EGFR L858R/WT
Guha08 exp1 delEGFR/WT
Guo08 MKN45 0hr / 3hr
Huang07 EGFRvIII M/DK
Kratchmarova05 EGF and PDGF / control
Guha08 exp1 mutRAS/WT
Bose06 Her2+Inhibitor/WT
B
EGFR
ERBB2
PIK3R1
CTNND1
PLCG1
BCAR1
MAPK3
SHC1
MAPK1
MAPK14
EPHA2
PTRF
INPPL1
ACP1
ANXA2
PTK2
CAV1
PXN

Figure 4. Bayesian network modeling of phosphorylation data from the current study and eight other published ERBB signaling related tyrosine phosphorylation data. Bayesian network modeling of phosphorylation data from the current study and eight other published ERBB signaling related tyrosine phosphorylation data. (A) The heatmap of the 18 proteins with the approximations for the missing data generated by the "nearest neighbor method" using a discrete distance measure. (B) The top scoring Bayesian network generated from the above phosphorylation data. The nodes are obtained from the phosphorylation ratios from the datasets. The edges represent direct and indirect causal influence. The arrows indicate the direction of causality. Nodes connected by black edges have the same phosphorylation level more often than not, indicating a positive influence. EGFR and ERBB2, were forced source nodes in the network; they were only allowed outgoing edges except between each other. Nodes were restricted to have no more than three parents. Yellow nodes are kinases and blue nodes phosphatases. Note central role of PTRF (Guha U., et al., PNAS 2008:105:14112-14117).

B. Study of the expression of all 48 nuclear hormone receptors (NRs) reveals tumor specific differences in microdissected tumor and normal lung epithelium. In collaboration with Dr. Wistuba, the Minna and Mangelsdorf laboratories at UTSW studied the expression of all 48 NRs using qRT PCR technology on RNA isolated from microdissected tumors and normal lung epithelium from 30 patients (Fig. 8). This approach revealed NRs whose expression differed between tumor and normal lung from the same patient, and also showed great variation in the expression of NRs between different lung cancers. Our ultimate goal would be to use these NRs expression profiles to hormonally manipulate lung cancer or as chemoprevention targets. We are interested if the pattern of NR expression in tumor may yield prognostic information on patient survival (Fig. 9). This approach was validated on an entirely separate lung cancer consortium data set of 442 lung adenocarcinomas, and demonstrates the biologic relevance of NR expression patterns in lung cancer.

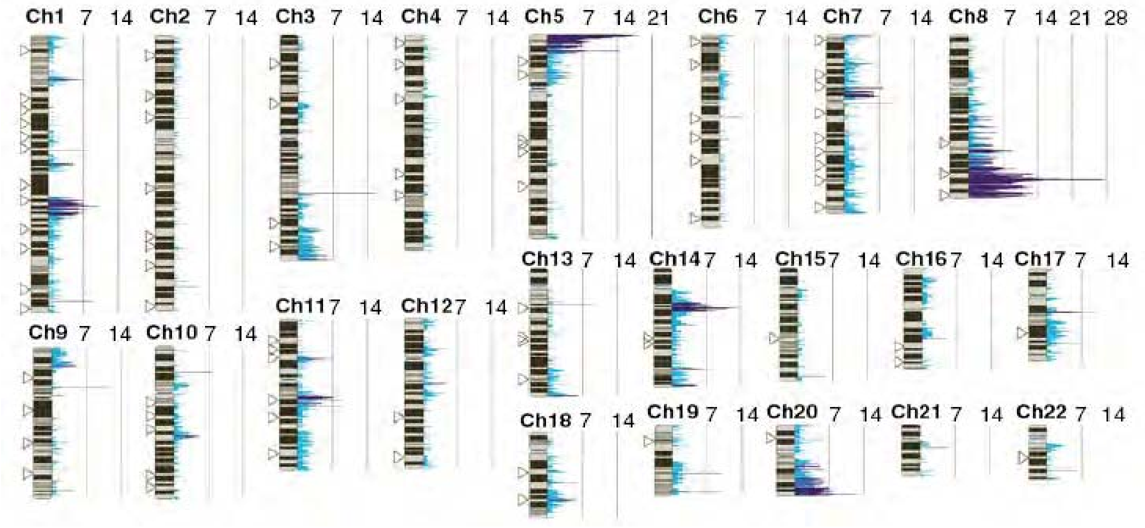


Figure 5. Hotspots of amplification present in cancer genomes. A histogram summarizing the regions of amplification detected on SMRT array v.2 (BAC array) across all 104 cancer cell line samples with the resulting values scaled to the segment with the highest count (shown as 28) and plotted against their corresponding genomic position. Hotspots are denoted by the dark blue shading, whereas the light blue shading represents regions amplified ~5 times. Triangles mark common fragile sites.
(Lockwood, Oncogene 27:4615, 2008).

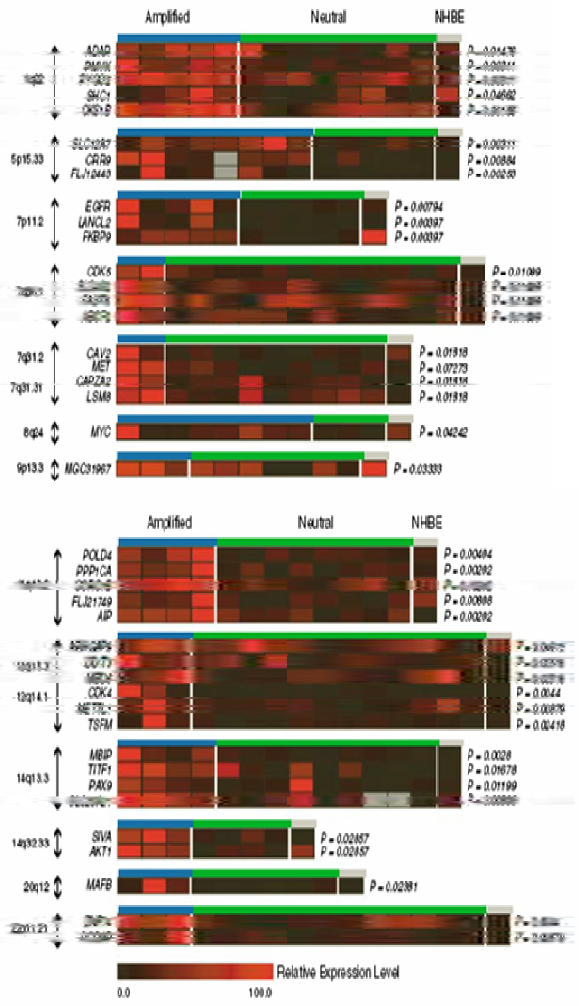


Figure 6. Impact of amplification on gene transcription levels. The relative expression values for genes with amplification and those with neutral copy number status are plotted as heatmaps for over-expressed genes from representative hotspots. The expression values for each gene have been normalized and scaled across the samples from 0 to 100. Red intensity indicates expression level. (Data from N = 27 NSCLC cell line samples and 3 HBEC samples using U133A and U133B Affymetrix array data with 23,583 unique genes). (From Lockwood, et al.,Oncogene 27:4615, 2008).

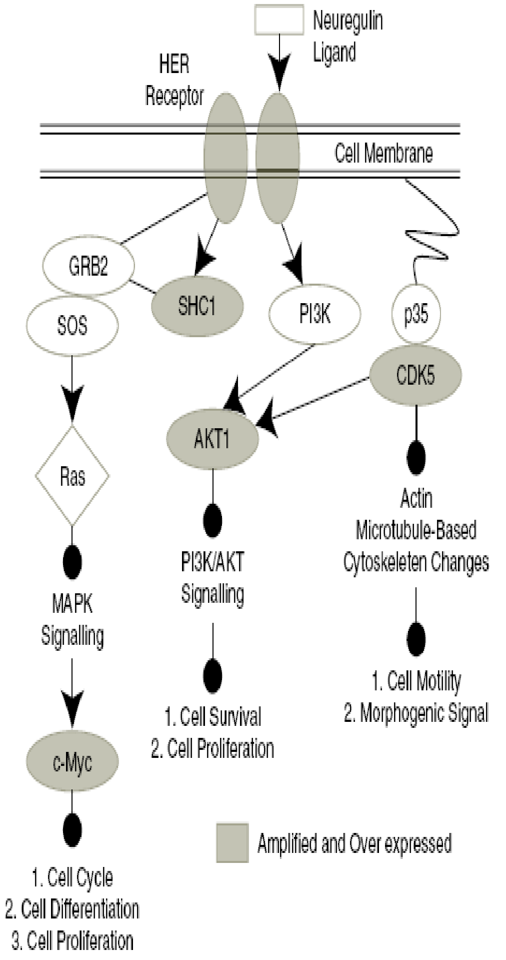

Neuregulin
Ligand
HER
Receptor
Cell Membrane
GRB2
SOS
SHC1
PI3K
p35
CDK5
AKT1
Ras
MAPK
Signalling
PI3K/AKT
Signalling
Actin
Microtubule-Based
Cytoskeleten Changes
1. Cell Survival
2. Cell Proliferation
1. Cell Motility
2. Morphogenic Signal
c-Myc
Amplified and Over expressed
1. Cell Cycle
2. Cell Differentiation
3. Cell Proliferation

Figure 7. Frequent amplification and overexpression of multiple EGFR-family-signaling components in non-small cell lung cancer (NSCLC). Diagram highlights the interaction of EGFR, SHC1, CDK5, SHC1 and MYC in the EGFR family-signaling pathway. Altered components are shaded grey and the effect on various cell pathways indicated (From Lockwood, et al., Oncogene 27:4615, 2008).

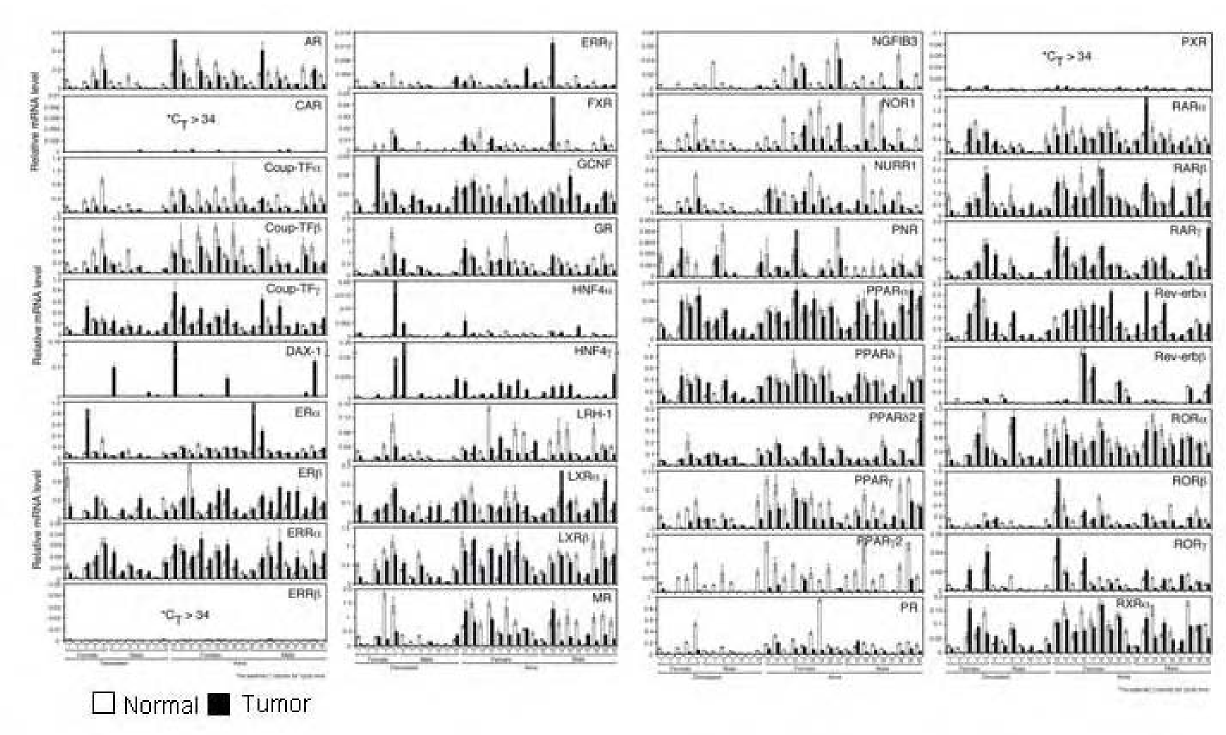


Figure 8. Quantitative mRNA expression of all 48 nuclear hormone receptors (NRs) in pair matched tumor and lung epithelium microdissected from the same patients (N = 30). Note differences in expression between tumor and normal epithelium in many patients for several receptors and the great variation in expression of NRs between different tumor samples. X axis, each of the 30 samples; Y axis, the level of mRNA expression for each receptor determined by Q RT PCR. White bars, expression of NRs in histologically normal lung epithelium; black bars, expression of the NR in the matched tumor sample.

Aim 4 Develop techniques to assess these molecular signatures in tissue specimens and serum obtained in Project 1, and assess the relevance of these molecular signatures as *in vivo* biomarkers using baseline and post-treatment specimens.

Summary of Research Findings

A. Identify regions amplified in preneoplastic lung lesions. The preferential selection of gene amplification events would drive tumorigenesis in specific cell lineages giving rise to lung cancer subtypes. As part of the aCGH studies in collaboration with Dr. Wan Lam, we discovered a new cell lineage specific genetic event that may provide novel target for new treatment strategies. Through integrative genetic analyses of multiple independent cohorts of clinical tumor samples, we identify the overexpression of *BRF2* (chromosome region 8p), a RNA polymerase III (Pol III) transcription initiation factor, as the result of increased gene dosage in the squamous cell lineage leading to lung squamous cell carcinoma (SqCC). We found that *BRF2* was amplified and overexpressed in preneoplastic carcinoma in situ (CIS) lesions in the airway epithelium of lung cancer patients (Fig. 10).

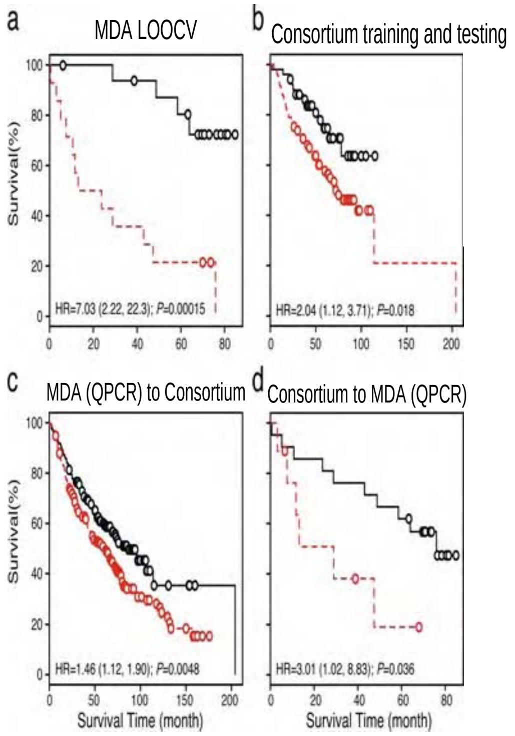


Figure 9. Expression of NRs in tumor provides prognostic information for the survival of non small cell lung cancer patients that is validated on a lung cancer consortium dataset. Prediction model of the NR signature is cross-validated by leave-one-out method, independently validated in multi-institutional consortium dataset, and cross-validated by different platforms of genetic analysis. (Multi-institutional data on N = 442 lung adenocarcinomas is from Shedden K, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. Nat Med 2008.) (LOOCV, leave-one-out cross validation). The MD Anderson Cancer Center (MDA) 30-patient data set was (a) first validated by LOOCV and then tested on the Consortium NR dataset derived from microarray expression profile data (panel C). Similarly, the Consortium data set was divided into training and testing components for NR expression (Panel B) and validated. Finally, the Consortium NR expression signature was validated on the MDACC data set (Panel D).

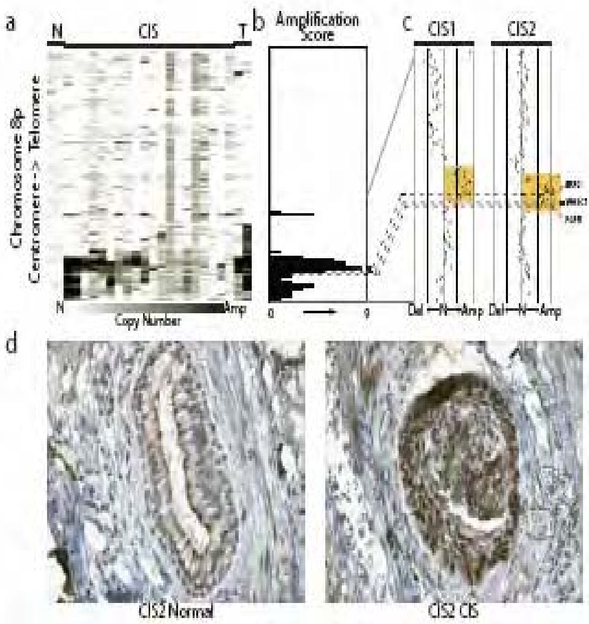


Figure 10. Amplification and overexpression of BRF2 in preinvasive SqCC lesions. (a) Frequent copy number increase of chromosome arm 8p in 20 bronchial CIS lesions. Samples are ordered in columns and ordered by genomic position along 8p. The color scale ranges from white (neutral copy number, N) to black (amplification, Amp). Data from representative normal lung (N) and SqCC tumor samples (T) are displayed to the left and right of the CIS cases respectively. (**b)** Amplification score along chromosome 8p for the 20 CIS cases. Regions of amplification were defined for each case and summarized across the group to determine the incidence of occurrence. Dashed lines represent the positions of *BRF2*, *WHSC1LC* and *FGFR1* from top to bottom respectively. (**c)** Array CGH copy number profiles for two individual CIS cases with 8p amplification. Each black dot represents an array element ordered by genomic position. Those shifted to the left of the middle line (N) have decreased copy number (Del) whereas those shifted to the right have increased copy number (Amp). Dashed lines represent the positions of the three genes as in **c.** The region highlighted in orange represents the region of high level amplification in each sample. The amplicon in CIS1 includes only *BRF2* with *WHSC1L1* and *FGFR1* outside or spanning the boundaries while the amplicon in CIS2 contains all three genes. (**d)** Immunostaining of CIS2 with anti-*BRF2* polyclonal antibody revealed elevated staining in CIS epithelia compared with normal from the same tissue section. (From Lockwood et al. BRF2 is a Lineage Specific Oncogene Amplified Early in Squamous Cell Lung Cancer Development. Proc Natl Acad Sci U S A 2008; (Submitted, PNAS 2008-11342.)

Frequent activation of BRF2 in pre-invasive bronchial carcinoma *in situ* and dysplastic lesions provides evidence that BRF2 expression is an early event in cancer development of this cell lineage. Ectopic expression of BRF2 in human bronchial epithelial cells induces a transformed phenotype and demonstrates downstream oncogenic effects, while siRNA-mediated knockdown suppresses growth of cells overexpressing BRF2 (data not shown). Our data suggests that genetic alteration of BRF2 represents a novel mechanism of lung tumorigenesis through the increase of Pol III-mediated transcription in cancer.

Aim 5. Identify gene expression signatures that characterize progression from immortalized to transformed to tumorigenic human bronchial epithelial cells based on already available high-throughput gene expression microarray data and validate these signatures using tissue microarrays (TMAs) containing normal bronchial epithelium, hyperplasia, squamous metaplasia, dysplasias, squamous cell carcinomas, atypical adenomatous hyperplasia, and adenocarcinomas (Years 4-5).

Summary of Research Findings

We used the *in vitro* human lung carcinogenesis model described in the previous year's report. This model includes normal human bronchial epithelial (NHBE) cells, BEAS-2B cells, which are NHBE immortalized with SV40 T/ Adeno12 virus and transformed (1198) and tumorigenic (1170-I) cells derived from BEAS-2B after exposure to cigarette smoke condensate *in vivo*. Immortalized HBECs (1799), derived from BEAS-2B without exposure to cigarette smoke condensate, were also used in this model. These isogenic cells represent normal, immortalized, transformed and tumorigenic cells and their study offers an opportunity to identify different progressive changes in genotype or epigenetic changes.

During last year, we further examined the functional significance and clinical relevance of the differentially expressed genes. We used the Ingenuity Pathways Analysis® software to gain a pathway level of understanding of genes differentially expressed between the normal, immortalized, transformed, and tumorigenic cells and how they fit within molecular interaction networks. We found that many of the differentially expressed genes (e.g., UBE2C, MCM2, FEN1, and BIRC5) function in cell cycle control and display significant interactions with each other and with other known cancer-related genes in significant topological gene networks (Fig. 11).

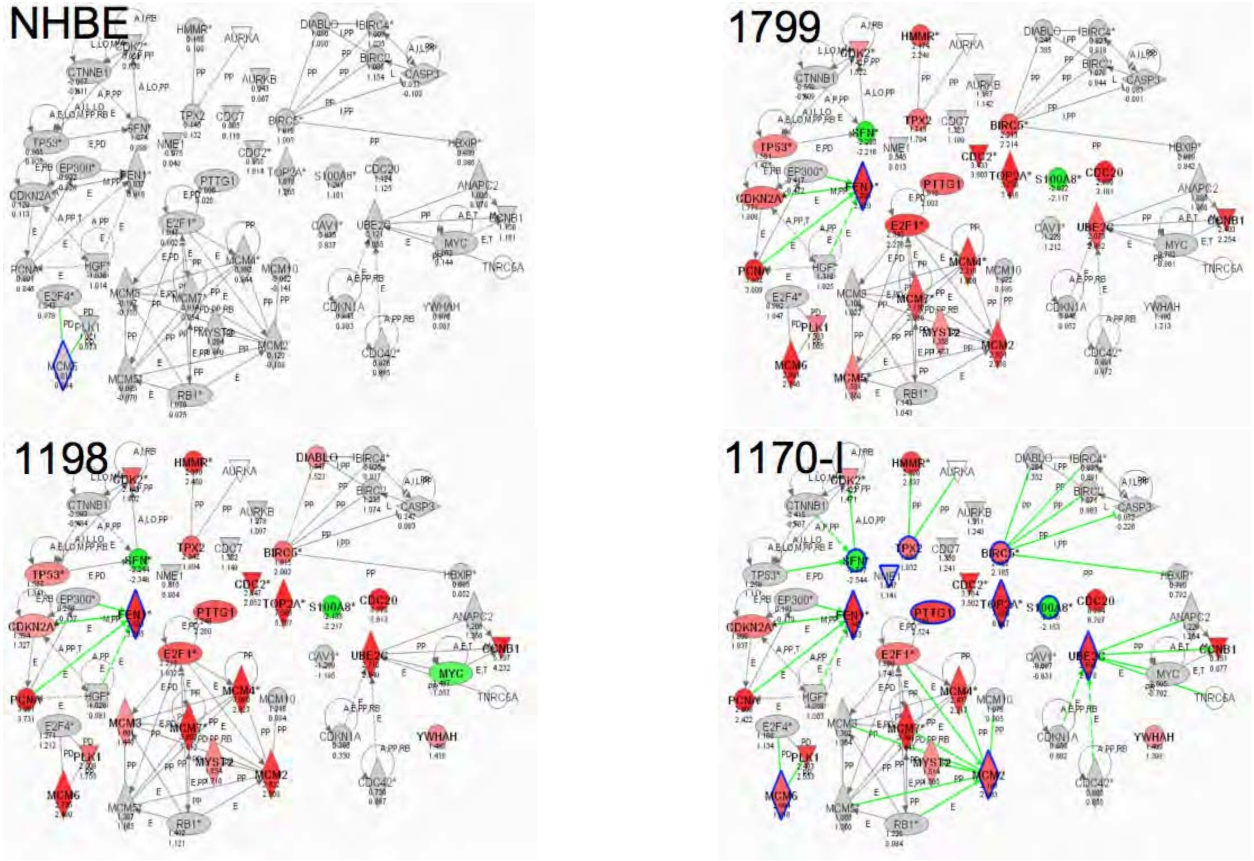

NHBE
1799
1198
1170-I
CDK2*
HMMR*
AURKA
DIABLO
BIRC4*
CTNNB1
AURKB
BIRC2
CASP3
TPX2
CDC7
BIRC5*
SFN*
NME1
CDC2*
HBXIP*
TP53*
EP300*
TOP2A*
S100A8*
CDC20
FEN1*
PTTG1
ANAPC2
CDKN2A*
CCNB1
E2F1*
CAV1*
UBE2C
MYC
PCNA*
HGF*
MCM4*
MCM10
TNRC6A
E2F4*
MCM3
MCM7*
CDKN1A
YWHAH
PLK1
MYST2
MCM2
MCM6
MCM5*
RB1*
CDC42*

Figure 11. Network of genes related to cell cycle progression significantly modulated in lung carcinogenesis model including normal (NHBE), B) immortalized (1799); C) transformed (1198); and D) tumorigenic (1170-I) cells. The changes in gene expression relative to normal NHBE (A, grey) are colored as follows: Red, upregulated; green, downregulated; grey, unchanged.

We validated several of the more differentially expressed genes by quantitative PCR and western blotting and found the expression of MCM2 (minichromosome maintenance 2), TPX2 [microtubule-associated, homolog (Xenopus laevis)], FEN (Flap endonuclease 1), and UBE2C (Ubiquitin conjugating enzyme E2C) increased from normal to tumorigenic states, whereas the expression of SFN (stratifin or 14-3-3-sigma) decreased (Fig 12). The expression of these genes in cell lines derived from human lung tumors was similar to the 1170 tumorigenic cells in our cell model.

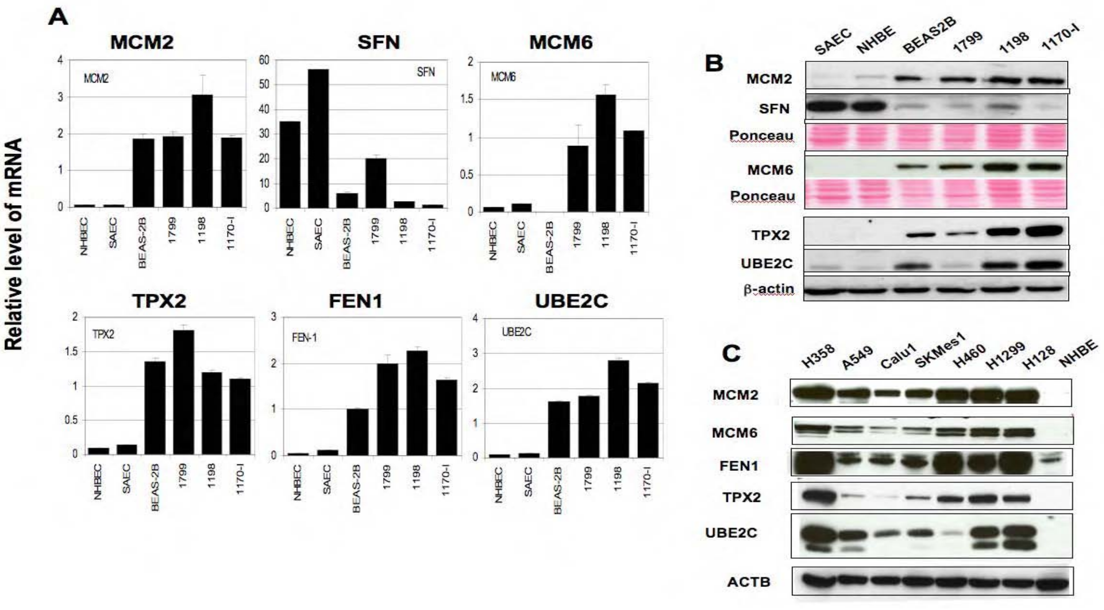


Figure 12. Validation of the differential expression of MCM2, SFN, TPX2, FEN, and UBE2C genes in cells comprising the in vitro carcinogenesis model (A) by quantitative PCR. Western blotting analysis of these selected genes in the in vitro carcinogenesis model (B) and several NSCLC cell lines from human tumors (C).

To determine the lung cancer relevance of the data obtained using the in vitro lung carcinogenesis system, we used *in silico* analysis of data published by Garber et al. and found that UBE2C expression was higher in lung tumors, both adenocarcinomas and squamous cell carcinomas, relative to normal lung tissue (Fig. 13).

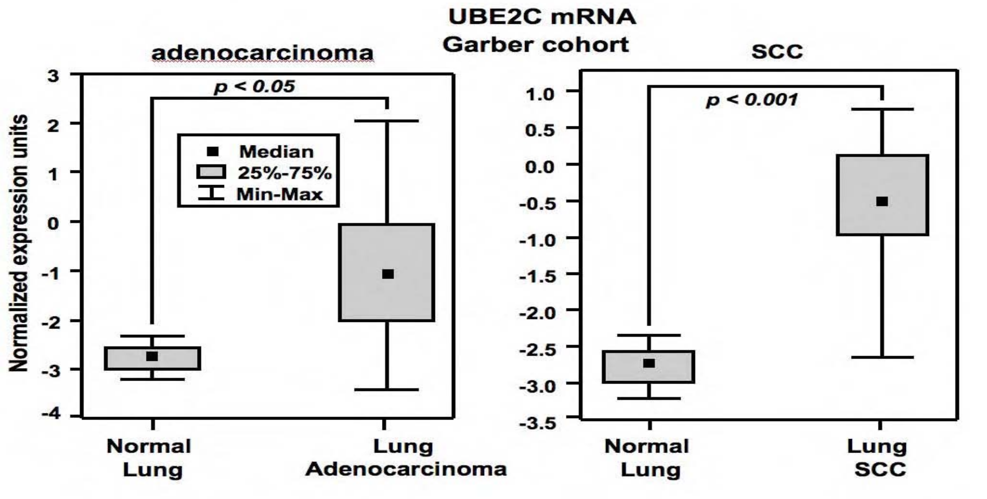


Figure 13. Validation of the differential expression of the UBE2C gene by *in silico* analysis of publicly available data of gene expression in lung cancer and normal lung tissue specimens from the Garber cohort.

Further studies were focused on UBE2C because it plays important roles in the ubiquitination pathway, which has been shown to affect cyclin-A degradation and control the process by which cells proceed from DNA segregation and cell division to a new round of DNA replication. UBE2C is over expressed in different types of cancer.

However, the relationship between UBE2C expression and lung cancer development and prognosis is not known. In collaboration with Dr. Ignacio Wistuba, we found using immunohistochemical methods that UBE2C protein level is undetectable in normal appearing lung epithelium, but its level increases in aberrantly proliferating lesions and premalignant lesions with a further increase with the severity of the progression to malignancy along the SCC development pathway (Fig. 14). In addition, there was an increased expression in adenocarcinoma compared to normal epithelium (Fig. 15).

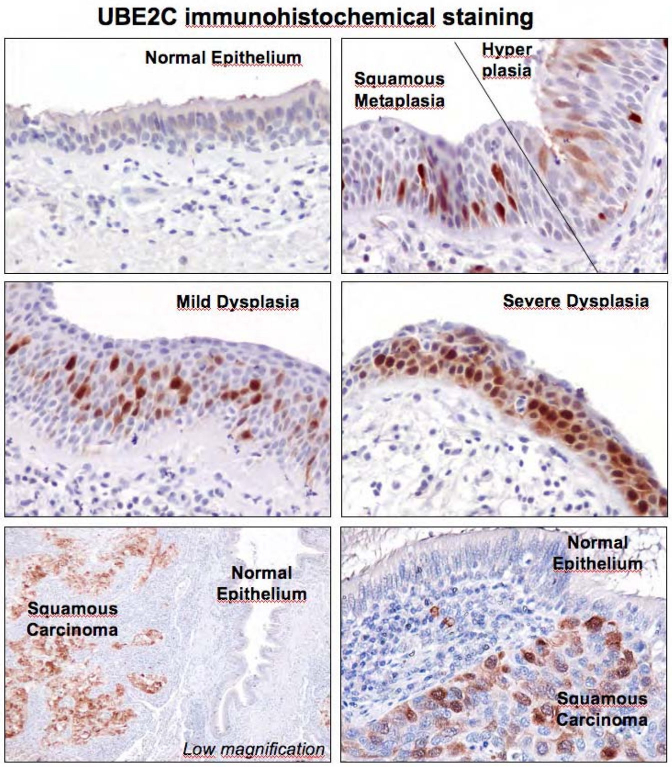

UBE2C immunohistochemical staining
Normal Epithelium
Squamous Metaplasia
Hyper plasia
Mild Dysplasia
Severe Dysplasia
Squamous Carcinoma
Normal Epithelium
Low magnification
Normal Epithelium
Squamous Carcinoma

Figure 14. Immunohistochemical staining of UBE2C in histological sections of lung tissues including normal epithelium, hyperplasia, squamous metaplasia, mild dysplasia, severe dysplasia, and squamous cell carcinoma (SCC).

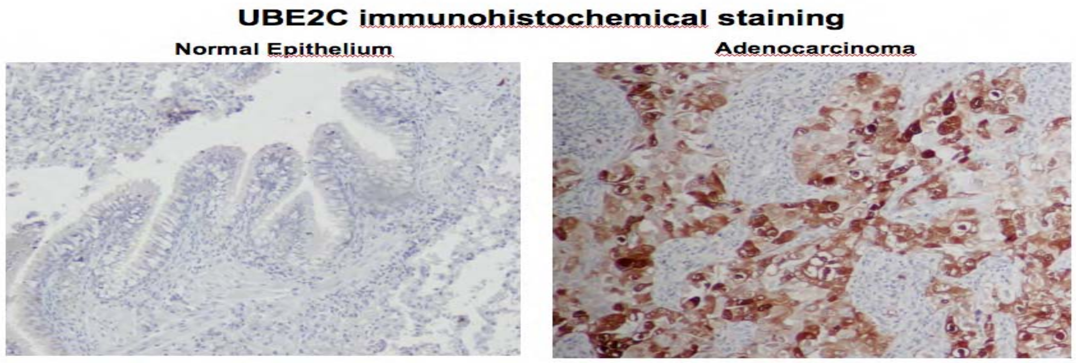

UBE2C immunohistochemical staining
Normal Epithelium
Adenocarcinoma

Figure 15. Immunohistochemical staining of UBE2C in histological sections of normal lung and lung adenocarcinoma.

Finally, we examined whether the expression of UBE2C is related to prognosis of lung cancer patients with stage I NSCLC. Figure 16 shows that both overall survival and progression-free survival were statistically significantly lower in those cases where UBE2C expression was positive.

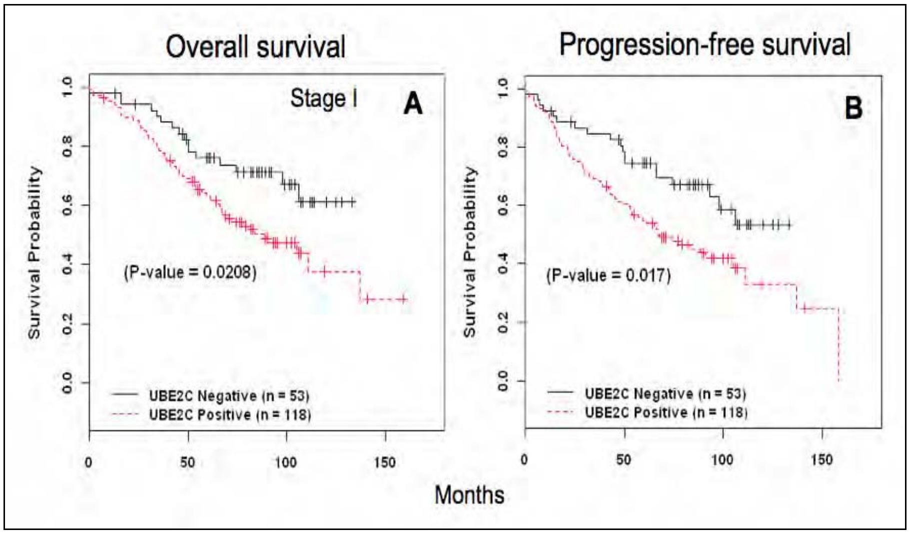


Figure 16. Kaplan-Meier plots and log-rank statistics for the overall survival (A) and progression-free survival (B) of stage I NSCLC patients separated according to UBE2C expression.

Aim 6. Identify gene expression signatures in bronchial brush specimens from the 50-60 patients enrolled in the Vanguard study using high-throughput genomics approach (Years 4-5).

Summary of Research Findings

1. Samples from patients enrolled in Vanguard project.

We have been collecting bronchial brush samples from Vanguard project (ID: 2003-0424). As of 12/12/08, we have collected samples from 45 patients at baseline, 29 patients at 12 months, 15 patients at 24 months, and 5 patients at 36 months during follow-up period (total 95 time points). A summary of the sample information is presented in the following table.

		Site of bronchial brush					
	No. of timepoints	**LB10***	**LUL***	**MC***	**RB10***	**RML***	**RUL***
Baseline	45	44	44	44	44	44	44
12 months	29	29	29	29	25	29	29
24 months	15	15	15	15	15	15	15
36 months	5	5	5	5	5	5	5
Total	95						560

* LB10, left lower lobe; LUL, left upper lobe; MC, main carina; RB10, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

2. RNA preparation

Total RNA have been extracted from all samples from 205 samples including all samples from MC sites.

3. Global gene expression analysis

Among the 205 RNA samples, we have successfully amplified and hybridized 138 samples using Affymetrix ST1.0 Human gene chips. The priority has been given to samples at the MC site collected at baseline and 12 months in order to address whether gene expression profiles at the main airway can be affected by lung cancer risk (comparing these profiles with those obtained from profiles of those without lung cancer) as well as by the presence of tumors (comparing the profiles with profiles obtained from lung cancer patients before tumors being resected). The following table summarizes samples with available gene expression profiles.

	No. of MC sites	**No. of sample (other site of diseased lung)**	**No. of sample (normal lung)**	**Total sites**
	63	37	38	138
No. of Patient samples	63	37		

We are now in the process to perform gene expression profiles for additional samples and will analyze the profiles to address the questions proposed in the Revitalization Proposal.

Key Research Accomplishments

- Generated HBECS from over 45 different individuals and immortalized 10 peripheral small airway epithelial cells (HSAECs).
- Testing “preneoplastic” HSAECs with chemoprevention agents and performing microarray mRNA expression profiles and array-based CGH copy number profiles.
- Compared 104 cancer lines representing diverse tissue origins and discovered an unexpected frequency of activated genes in amplification 'hotspots.'
- Using qRT PCR technology on RNA isolated from microdissected tumors and normal lung epithelium from 30 patients, found nuclear hormone receptors whose expression differed between tumor and normal lung from the same patient, and between different lung cancers.
- Discovered that *BRF2* was amplified and overexpressed in preneoplastic carcinoma in situ (CIS) lesions in the airway epithelium of lung cancer patients
- Differentially expressed genes such as UBE2C, MCM2, FEN1, and BIRC5 are located within gene interaction networks significantly modulated between normal, immortalized, transformed and tumorigenic cells.
- The findings suggest that changes in MCM2, SFN, TPX2, FEN, and UBE2C genes are all important for the replication of immortalized, transformed and tumorigenic cells.
- UBE2C protein level is undetectable in normal appearing lung epithelium, but its level increases in aberrantly proliferating lesions and premalignant lesions, with a further increase with the severity of the progression to malignancy along the SCC development pathway. In addition, there was an increased expression in adenocarcinoma compared to normal epithelium
- In lung cancer patients with stage I NSCLC, both overall survival and progression-free survival were statistically significantly lower in those cases where UBE2C expression was positive.
- Successfully amplified and hybridized 138 samples using Affymetrix ST1.0 Human gene chips.

Conclusions

- Amplification is a far more common mechanism of oncogene activation than previously believed and specific regions of the genome are hotspots of amplification. These amplicons provide important new biomarkers that may be present in preneoplastic lesions and could be promising targets for the monitoring and chemoprevention of lung cancer.
- NRs may provide prognostic value in lung cancer patients.
- Genetic alteration of BRF2 represents a novel mechanism of lung tumorigenesis through the increase of Pol III-mediated transcription in cancer.
- Differentially expressed genes (e.g., UBE2C, MCM2, FEN1, and BIRC5) function in cell cycle control and display significant interactions with each other and with other known cancer-related genes in significant topological gene networks.
- UBE2C appears to be an promising prognostic biomarker for lung cancer survival.
- Patient survival and progression-free survival were statistically significantly lower in those cases where UBE2C expression was positive.

Project 3: Premalignant Bronchial Epithelia: Molecular and Cellular Characterization of Lung Tumorigenesis

(PI and co-PIs: Walter Hittelman, Ph.D., Ja Seok Koo, Ph.D., Rodolfo C. Morice, M.D.)

Aim 1 Identify and characterize differentially expressed genes in the LIFE bronchoscopy-identified abnormal areas of the bronchial epithelia of of enrolled subjects in VITAL trials.

Previous studies have shown that bronchial regions that appear abnormal by light-induced fluorescence endoscopy (LIFE) bronchoscopy show increased genetic changes when compared to normal-appearing sites, even if there are no differences in histological appearance. Since LIFE-positive lesions are at increased risk for cancer development, especially when they contain particular genetic alterations, we hypothesize that these LIFE-positive sites represent lesions at an early stage of tumorigenesis and may differentially express genes important for driving tumorigenesis. Thus, comparative gene expression analyses between LIFE-positive and LIFE-negative sites within the same individual may provide a filter for identifying genes whose levels of expression are important for driving tumorigenesis.

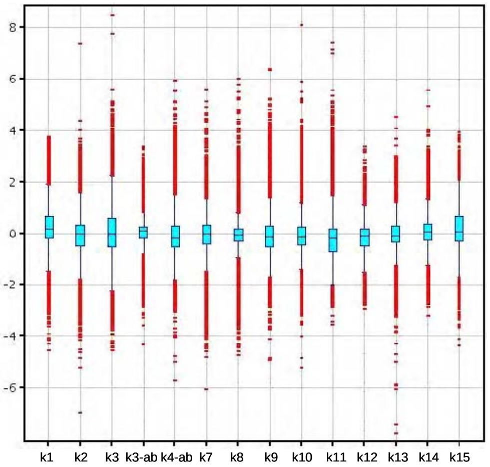


Figure 1. The distribution of gene expression ratio of abnormal area vs. normal area among all brush samples. RNA prepared from abnormal area was labeled with Cy5 and normal area with Cy3. The two sources of RNA then mixed and hybridized with Agilent Whole Human Genome Oligo Microarray. The measured intensity ratio between the two fluorphores thus represents the relative expression level of each gene. Box and whisker plots were generated with GeneSpring program.

Summary of Research Findings

Whole genome expression profiling microarray analysis produces valuable information that was not previously available, but is noted for its difficulty in obtaining consensus results in sample analysis. To minimize the sample variation, we used paired bronchial brush samples from the normal and abnormal areas of each individual to perform microarray analysis. To date, a total of 14 sample pairs have been submitted to the MDACC microarray core facility for analysis. The majority of sample pairs are from “white-light abnormal” area vs. “white-light normal” area, except cases k3 and k4, which are from LIFE-abnormal while white-light normal area. Nevertheless, there is no obvious disparity among these samples in terms of their general gene expression levels (Fig. 1). To analyze these microarray datasets, we first examined the commonly regulated genes using the expression level of all the sample sets. Sorting the genes by their average expression level further reveals the heterogeneity of the samples population and revealed the presence of possible false- positive results in the sample set (Fig. 2). To determine which sub-set of samples might be falsely positive or grossly different in their genetic makeup, we performed an unsupervised classification using a complete-linkage hierarchical clustering analysis. The analysis distinctively divided the samples into two sample clusters, which we designated as C1 and C2 (Fig. 3).

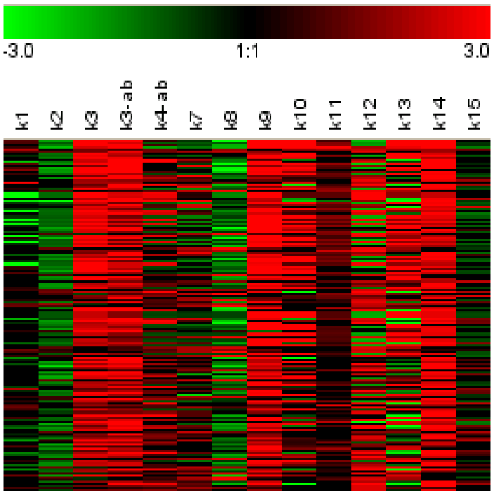


Figure 2. Heterogeneity in the genome expression profile of samples. Only gene expression levels with significant signal ratio ($p<0.05$) in at least 50% of the microarray samples are included (total 4816 probes for 3867 genes). The genes with expression levels that did not reach significance are designated as 0, for no change. The genes were sorted in the descending order of average expression level. The top 100 genes are shown. Heat map was generated with Genesis program.

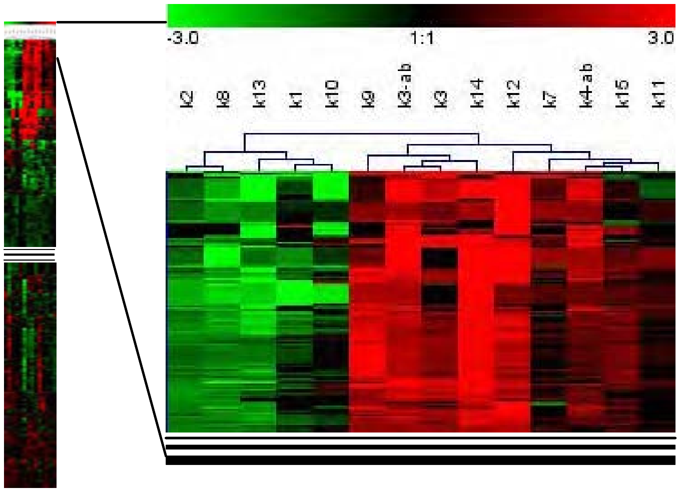


Fig 3. Hierarchical clustering analysis divided all the brush samples into two distinctive sample clusters. The probe set meets the criteria mentioned in Fig 2 was used to perform compete linkage hierarchical analysis. Left: portions of the heat map were shown; right: top portion of heat map was enlarged to show the sample label.

We again compared these two clusters by examining their top up- and down-regulated genes separately. However, these top regulated genes still show significant heterogeneity within each cluster (Fig. 4) and they further suggested that our measured results can be easily affected by some outliers with unusual expression levels, such as several hemoglobin genes. The presence of outliers in the sample set may be caused by the different amount of blood retained in the brush samples.

To avoid such bias, we ran a K-means clustering analysis within each sample cluster, and set the cluster number to 50 to obtain an average of about 100 gene probes per cluster (Figs. 5 and 6). We predicted that the important differentially regulated genes should be consistently up or down within each sample cluster, and we selected only the gene clusters that are constantly up- or down-regulated across the samples (i.e., gene clusters 28, 29 and 40 in C1; gene cluster 1, 10, 11, 46 and 47 in C2) for further analysis as genes of interest. As shown in Table 1 and 2, these genes of interest are disparate between the two sample clusters with slight overlap and an inverse correlation in certain gene clusters. These results indicate that either the patients from these two clusters have a distinct genetic make-up or samples from one of the clusters are simply false positive, providing a practical contrast for the other.

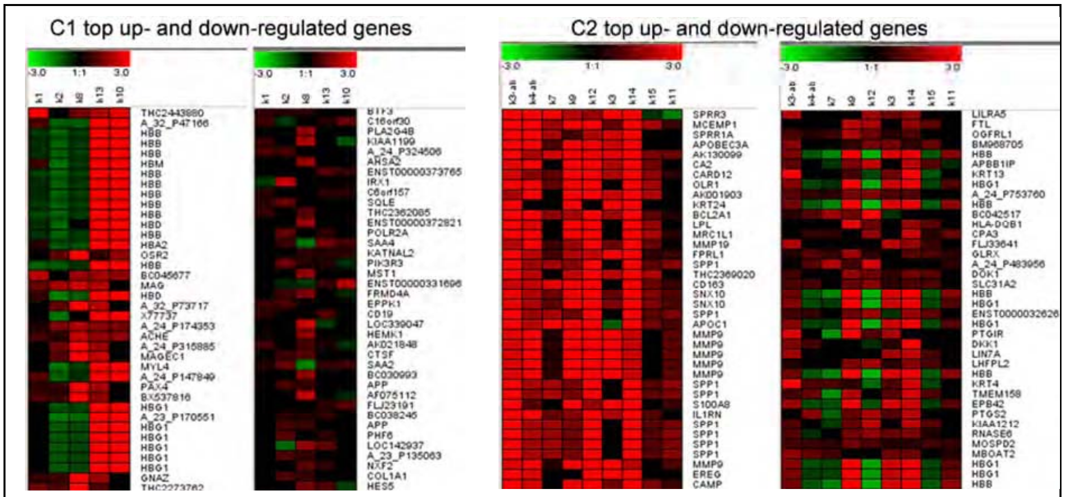


Figure 4. Top up- and down-regulated genes in the two sample clusters. The gene expression ratios in each sample cluster were averaged for each gene, and then sorted in descending order

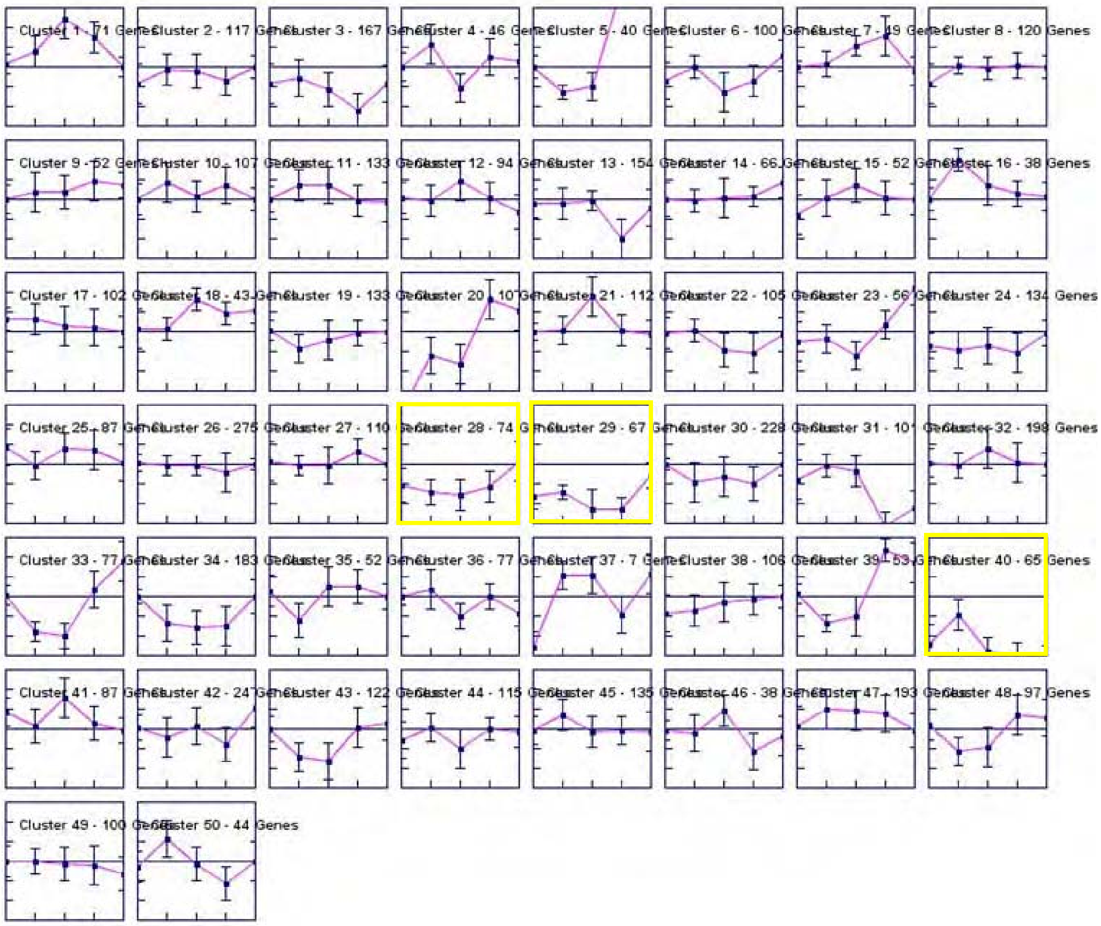


Figure 5. The centroid view of the 50 clusters resolved with k-means clustering analysis for sample cluster C1. Gene clusters 28, 29 and 40, which show consistent change (down-regulated) across the samples were chosen (highlighted).

As the selected gene clusters in C2 consist of several genes that were previously reported to be relevant to tumorigenesis, we extracted these genes and analyzed the potential key signal pathways that are involved with a Kegg pathway analysis program. The pathway analysis reveals that the cytokine-cytokine receptor pathway is the top signal pathway involved, followed by other tumorigenesis pathways, such as cell migration and adhesion (Table 3). Due to our previous suspicions regarding false-positive C1 samples, we used these two clusters to run a differential expression analysis with GEPAS (Gene Expression Pattern Analysis Suite) program (Fig. 7). Genes in the top and bottom list encompass numerous genes that were previously reported to be important for tumorigenesis. Results further indicate that the two sample clusters represent two distinct populations; thus, the two sample sets need to be analyzed separately to uncover additional information useful to future studies. Microarray analysis of the sample clusters may provide important information for identifying the molecular differences between the abnormal and normal tissue at early stage of tumorigenesis in future studies.

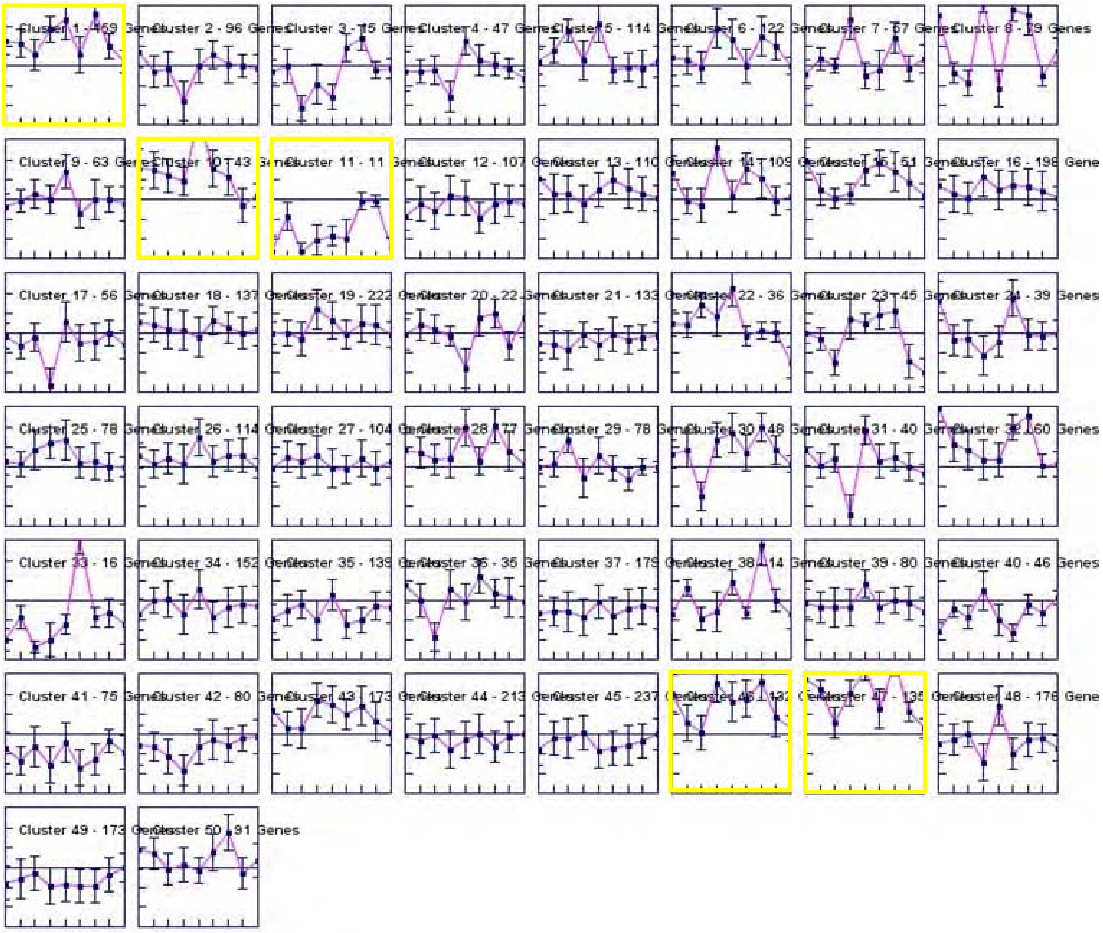


Figure 6. The centroid view of the 50 clusters resolved with k-means clustering analysis for sample cluster C2. Gene clusters 1, 10, 46, 47 (up-regulated), and 11 (down-regulated) which show consistent change across the samples were chosen (highlighted).

In conclusion, whole genome expression microarray analysis provides tremendous information about the molecular differences between abnormal and normal tissues. However, the real signal that dictates the pathological event may be muffled by the noises from mixed sample population. By distinguishing the subsets of the sample population, we were able to identify the consistently regulated genes and the signaling pathway that is critical for tumorigenesis.

C1 Cluster28	FoldChange	C1 Cluster29	FoldChange	C1 Cluster40	FoldChange
ANPEP	0.418	AIF1	0.368	APOC1	0.139
AOAH	0.549	AQP9	0.377	APOE	0.160
AP1S2	0.562	BCL2A1	0.278	BCAT1	0.254
APOBEC3A	0.430	CA2	0.321	CCL18	0.171
CD14	0.417	CAMP	0.344	CCL2	0.147
CD1C	0.511	CARD12	0.314	CCL23	0.197
CD300A	0.561	CCL3L3	0.271	CD163	0.187
CDR1	0.505	CCR1	0.306	CPM	0.227
CLEC4E	0.554	CD300LF	0.324	CXCL5	0.214
CSF2RA	0.454	CD33	0.348	FN1	0.146
DOCK7	0.457	CD36	0.276	HAVCR2	0.264
DPEP2	0.480	CD52	0.323	HTRA4	0.221
EMR2	0.521	CHIT1	0.212	IL1B	0.214
EVI2A	0.489	CPM	0.227	KRT24	0.320
FCGR3A	0.512	CYBB	0.215	LPL	0.151
FCGR3B	0.565	FPRL2	0.280	MARCO	0.223
FES	0.508	GPR65	0.360	MMP12	0.090
FGR	0.425	HCK	0.386	MMP19	0.182
FPR1	0.524	HK2	0.373	MRC1L1	0.242
FYB	0.471	HK3	0.286	MS4A4A	0.195
GMFG	0.506	HLA-DQA1	0.379	MS4A7	0.267
HTLF	0.505	LAIR1	0.285	MSR1	0.184
IFIT2	0.418	LAPTM5	0.318	OSM	0.236
IGSF6	0.447	LILRB3	0.378	PLA2G7	0.200
IL8RA	0.588	LY96	0.338	SPRR3	0.171
ITGAM	0.591	MYO1G	0.347	VSIG4	0.219
ITGAX	0.393	NCKAP1L	0.363		
ITGB2	0.516	OASL	0.361		
LAT2	0.457	PARVB	0.361		
LILRA4	0.537	PIK3R5	0.367		
LILRB2	0.521	PLAUR	0.329		
LRRC25	0.434	PLEK	0.347		
MPP1	0.511	SDS	0.246		
MS4A6A	0.480	SIGLEC7	0.281		
MTPN	0.577	SLC8A1	0.417		
MYL9	0.494	SPARC	0.360		
NCF1	0.482	SPP1	0.260		
P2RY14	0.518	TBXAS1	0.357		
PRG1	0.444	TNFAIP6	0.258		
PSCD4	0.528	TYROBP	0.371		
PTPN7	0.546	URP2	0.264		
RAB32	0.461				
RIN3	0.506				
SH2B3	0.413				
SIGLEC9	0.412				
SNAI3	0.577				
TGFB1	0.541				
TLR4	0.404				
TMEM140	0.555				
TNFAIP8L2	0.470				
TNFSF13B	0.496				
TRPM2	0.481				
TSPAN33	0.476				
UNC13D	0.389				
VASP	0.442				

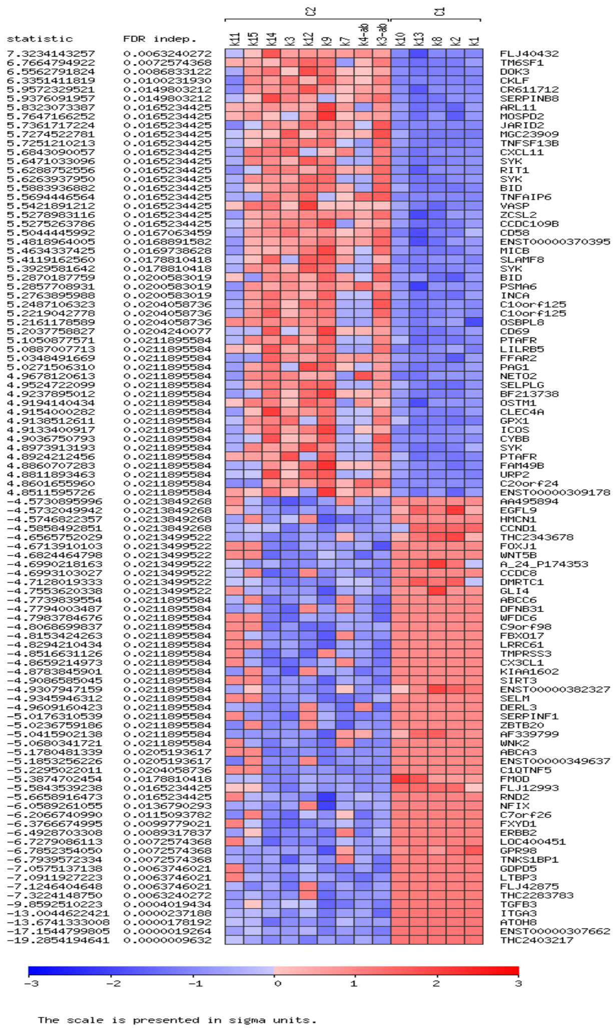
Table 1. The gene clusters showing consistent regulation of expression in sample cluster C1. Fold change is the average relative expression level of all samples in C1.

Cluster1	F-C	Cluster10	F-C	Cluster11	F-C	Cluster46	F-C	Cluster47	F-C
ACP5	2.41	CCL2	2.17	ACTG2	0.33	AATK	2.94	ADAMTSL4	3.19
ADCY7	3.54	CCL3	2.25	COL6A2	0.38	AIF1	2.27	ADCY7	3.54
ADFP	2.74	CCL4	2.25	DES	0.21	AMICA1	2.12	AGRP	3.50
ADM	2.69	CCRL2	3.16	FLJ30901	0.47	AOAH	2.09	ALOX5AP	3.53
ALOX5	2.72	CXCL5	3.78	GREM2	0.50	AP1S2	3.09	AOC3	2.76
APOC2	2.40	CXCR4	3.26	MFAP4	0.40	APBB1IP	2.36	APOBEC3A	5.82
APOE	2.76	FCAR	1.72	MYH11	0.24	APOBEC3B	2.67	APOC1	4.48
C1QA	1.96	G0S2	1.55	PDLIM3	0.29	ARHGAP15	2.30	AQP9	4.19
C1QB	2.25	IL1A	3.08	TBX2	0.28	ARL11	2.22	BCAT1	3.70
C1QC	2.17	IL8	1.58	WISP2	0.35	BIN2	2.07	BCL2A1	4.84
C8A	3.05	MMP12	3.06			BTK	3.20	CA2	5.06
CATSPER1	3.07	ORM1	1.59			CCR7	2.19	CAMP	4.20
CCL3L3	2.51	OSM	2.36			CD300LF	2.36	CARD12	5.04
CCRL2	3.16	PHACTR1	2.12			CD33	2.78	CCL18	3.74
CD109	3.15	PLA2G7	3.09			CD48	2.40	CCL23	3.76
CD274	2.64	SDS	2.36			CD53	3.14	CCR1	3.20
CD300C	2.02	SLC26A4	1.74			CLEC4A	2.57	CD163	4.56
CD44	2.06	SPINK1	2.34			COTL1	2.22	CD36	3.10
CD68	2.54	SPRR2C	2.86			CPM	3.62	CD52	3.07
CD69	2.49	TKTL1	2.47			CPVL	2.29	CLEC12A	2.99
CD83	2.48					CREB5	2.66	CLEC4D	3.53
CD86	3.17					CRTAM	3.21	CLEC4E	4.12
CPM	3.62					CSF2RA	2.42	COLEC12	4.07
CXCL5	3.78					CXCL11	3.97	EBI2	2.77
DAB2	2.39					CYBB	3.47	EREG	4.21
DOCK11	2.93					DPEP2	2.02	FABP4	3.98
DOK2	2.24					EMR1	3.61	FGR	3.99
EMILIN2	1.96					EVI2B	2.39	FN1	2.65
EMP1	2.26					FCGR2A	2.89	FPRL1	3.48
FABP5	2.13					FCGR3A	4.00	GPA33	4.62
FAM89A	2.65					FCGR3B	3.28	GPNMB	3.84
FBP1	2.14					FCN1	2.29	HK3	4.10
FCER1G	2.38					FLI1	2.45	HP	3.04
FGR	3.99					FPR1	2.89	HPR	3.25
FLJ25416	2.40					FYB	2.46	HTRA4	3.45
HAVCR2	2.98					GLIPR1	2.49	IL1A	2.99
HMOX1	2.72					GLT1D1	2.71	IL1B	3.08
IFI30	2.88					GMFG	2.77	IL1RN	3.51
IL10RB	2.60					GNG2	2.20	INHBA	4.36
IRF8	3.02					GPR34	2.16	KIAA1212	3.60
ITGAX	2.37					GPR65	3.32	LPL	2.95
JAKMIP2	2.67					HCK	2.58	MARCO	4.72
KCNAB1	2.68					HK3	3.04	MCEMP1	3.83
KIAA1212	2.06					IGSF6	3.31	MMP19	6.19
KMO	2.50					IL8RA	1.94	MMP9	4.69
LAIR1	2.48					IL8RB	2.31	MRC1L1	4.70
LAPTM5	2.83					ITGAM	3.24	MS4A4A	4.07
LILRB3	2.90					ITGB2	3.30	MS4A7	3.27
LY86	2.56					LAPTM5	2.83	MSR1	3.82
MYO1G	2.05					LAT2	2.33	OLR1	4.96
NPL	2.38					LCP1	2.88	PCOLCE2	4.11
OASL	3.00					LILRA2	2.76	PLAUR	2.94
OSCAR	2.56					LILRB2	2.25	PLEK	3.24
PAG1	2.35					LRRC25	2.77	PPARG	3.35
PARVB	1.96					LST1	2.18	PRG1	2.94
PBEF1	2.16					LTB	2.13	PROK2	3.96
PHLDA1	2.10					LY96	2.82	RETN	3.09
PLAU	2.69					MFNG	2.09	SLA	3.70
PTPRO	2.75					MMD	2.54	SLC11A1	3.13
RAB8B	2.32					MME	3.47	SLC2A14	3.10
RBP4	2.68					MNDA	3.42	SNX10	4.55
SAMSN1	2.82					MPEG1	2.12	SPP1	4.58
SAPS1	2.32					MPP1	2.89	SPRR1A	5.84
SCD	2.45					MS4A6A	2.11	SPRR3	8.51
SLAMF8	2.43					NCF1	2.47	STAC	3.65
SLC11A1	3.13					NCF4	2.50	TEX14	2.98
SLCO2B1	3.09					PSCDBP	2.21	TM7SF4	3.32
TDRD9	2.00					PTCRA	2.29	TNFAIP6	4.13
THBD	1.97					PTGS1	2.08	TREM1	3.55
TIMP2	2.08					PTPRC	2.35	TUBB6	3.48
TLR4	2.34					QKI	2.68	VSIG4	3.44
TNF	2.31					RCSD1	2.35	W60781	4.11
TRPV2	2.39					RGS2	2.68	ZFHX1B	3.49
TYROBP	2.65					S100A8	4.38		
UPP1	2.87					S100A9	2.76		
VIM	2.29					SIGLEC5	2.51		
VMD2	2.24					SLC24A4	2.59		
						TAGAP	2.26		
						TBXAS1	2.15		
						TLR8	3.53		
						TREML2	2.45		
						TUBB3	2.21		
						WASPIP	2.36		

Table 2. The gene clusters showing consistent regulation of expression in sample cluster C2.

Pathway	**genes**
Cytokine-cytokine receptor interaction	CCR1, CCR7, CSF2RA, IL1A, IL1B, IL8, IL8RA, IL8RB, IL10RB, INHBA, LTB, OSM, CCL2, CCL3, CCL4, CCL18, CCL23, ALOX5, PTGS1, TBXAS1.
Toll-like receptor signaling pathway	LY96, IL1B, IL8, TLR8, CCL3, CCL4, CXCL11, SPP1, TLR4, TNFA, CD86
Leukocyte transendothelial migration	CYBB, ITGAM, ITGB2, MMP9, NCF4, NCF1, CXCR4
Cell adhesion molecules (CAMs)	CD274, ITGAM, ITGB2, PTPRC, CD86
ECM-receptor interaction	COL6A2, FN1, SPP1, CD36, CD44
Regulation of actin cytoskeleton	FN1, ITGAM, ITGAX, ITGB2
Focal adhesion	COL6A2, FN1, PARVB, SPP1
Arachidonic acid metabolism	ALOX5, PTGS1, TBXAS1

Table 3. The major canonical pathways in which the constantly up-regulated genes from sample cluster C2 are involved.



statistic	FDR indep.	
7.3234143257	0.0063240272	FLJ40432
6.7664794922	0.0072574368	TM6SF1
6.5562791824	0.0086833122	DOK3
6.3351411819	0.0100231930	CKLF
5.9572329521	0.0149803212	CR611712
5.9376091957	0.0149803212	SERPINB8
5.8323073387	0.0165234425	ARL11
5.7647166252	0.0165234425	MOSPD2
5.7361717224	0.0165234425	JARID2
5.7274522781	0.0165234425	MGC23909
5.7251210213	0.0165234425	TNFSF13B
5.6843090057	0.0165234425	CXCL11
5.6471033096	0.0165234425	SYK
5.6288752556	0.0165234425	RIT1
5.6263937950	0.0165234425	SYK
5.5883936882	0.0165234425	BID
5.5694446564	0.0165234425	TNFAIP6
5.5421891212	0.0165234425	VASP
5.5278983116	0.0165234425	ZCSL2
5.5275263786	0.0165234425	CCDC109B
5.5044445992	0.0167063459	CD58
5.4818964005	0.0168891582	ENST00000370395
5.4634337425	0.0169738628	MICB
5.4119162560	0.0178810418	SLAMF8
5.3929581642	0.0178810418	SYK
5.2870187759	0.0200583019	BID
5.2857708931	0.0200583019	PSMA6
5.2763895988	0.0200583019	INCA
5.2487106323	0.0204058736	C10orf125
5.2219042778	0.0204058736	C10orf125
5.2161178589	0.0204058736	OSBPL8
5.2037758827	0.0204240077	CD69
5.1050877571	0.0211895584	PTAFR
5.0887007713	0.0211895584	LILRB5
5.0348491669	0.0211895584	FFAR2
5.0271506310	0.0211895584	PAG1
4.9678120613	0.0211895584	NETO2
4.9524722099	0.0211895584	SELPLG
4.9237895012	0.0211895584	BF213738
4.9194140434	0.0211895584	OSTM1
4.9154000282	0.0211895584	CLEC4A
4.9138512611	0.0211895584	GPX1
4.9133400917	0.0211895584	ICOS
4.9036750793	0.0211895584	CYBB
4.8973913193	0.0211895584	SYK
4.8924212456	0.0211895584	PTAFR
4.8860707283	0.0211895584	FAM49B
4.8811893463	0.0211895584	URP2
4.8601655960	0.0211895584	C20orf24
4.8511595726	0.0211895584	ENST00000309178
-4.5730895996	0.0213849268	AA495894
-4.5732049942	0.0213849268	EGFL9
-4.5746822357	0.0213849268	HMCN1
-4.5858492851	0.0213849268	CCND1
-4.6565752029	0.0213499522	THC2343678
-4.6713910103	0.0213499522	FOXJ1
-4.6824464798	0.0213499522	WNT5B
-4.6990218163	0.0213499522	A_24_P174353
-4.6993103027	0.0213499522	CCDC8
-4.7128019333	0.0213499522	DMRTC1
-4.7553620338	0.0213499522	GLI4
-4.7739839554	0.0211895584	ABCC6
-4.7794003487	0.0211895584	DFNB31
-4.7983784676	0.0211895584	WFDC6
-4.8068699837	0.0211895584	C9orf98
-4.8153424263	0.0211895584	FBXO17
-4.8294210434	0.0211895584	LRRC61
-4.8516631126	0.0211895584	TMPRSS3
-4.8659214973	0.0211895584	CX3CL1
-4.8783845901	0.0211895584	KIAA1602
-4.9086585045	0.0211895584	SIRT3
-4.9307947159	0.0211895584	ENST00000382327
-4.9345946312	0.0211895584	SELM
-4.9609160423	0.0211895584	DERL3
-5.0176310539	0.0211895584	SERPINF1
-5.0236759186	0.0211895584	ZBTB20
-5.0415902138	0.0211895584	AF339799
-5.0680341721	0.0211895584	WNK2
-5.1780481339	0.0205193617	ABCA3
-5.1853256226	0.0205193617	ENST00000349637
-5.2295022011	0.0204058736	C1QTNF5
-5.3874702454	0.0178810418	FMOD
-5.5843539238	0.0165234425	FLJ12993
-5.6658916473	0.0165234425	RND2
-6.0589261055	0.0136790293	NFIX
-6.2066740990	0.0115093782	C7orf26
-6.3766674995	0.0099779021	FXYD1
-6.4928703308	0.0089317837	ERBB2
-6.7279086113	0.0072574368	LOC400451
-6.7852354050	0.0072574368	GPR98
-6.7939572334	0.0072574368	TNKS1BP1
-7.0575137138	0.0063746021	GDPD5
-7.0911927223	0.0063746021	LTBP3
-7.1246404648	0.0063746021	FLJ42875
-7.3224148750	0.0063240272	THC2283783
-9.8592510223	0.0004019434	TGFB3
-13.0044622421	0.0000237188	ITGA3
-13.6741333008	0.0000178192	ATOH8
-17.1544799805	0.0000019264	ENST00000307662
-19.2854194641	0.0000009632	THC2403217

The scale is presented in sigma units.

Figure 7. The top and bottom differentially expressed genes between the sample cluster C1 and C2. The analysis and heat map are performed and generated with GEPAS (Gene Expression Pattern Analysis Suite) program.

Aim 2 Establish an organotypic model system that mimics *in vivo* interactions between normal, premalignant, and malignant bronchial epithelial cells in the lung using cells derived from bronchial biopsies and immortalized bronchial cells.

Our prior studies using chromosome *in situ* hybridization to visualize genetic changes in the bronchial epithelium of current and former smokers suggested that, over years of tobacco smoke exposure, the combination of accumulating genetic damage, ongoing tissue damage, and wound healing results in a mosaic of evolving clonal outgrowths throughout the bronchial epithelium. To better understand the molecular basis of preferential outgrowth of more advanced bronchial epithelial clones, we proposed to utilize a cell culture model whereby normal and abnormal bronchial epithelial cells are grown on collagen or stromal cell-coated, suspended filters and exposed to an air-liquid interface. This organotypic culture environment mimics lung stratified epithelium, complete with basal cells, ciliated columnar cells, and mucus-producing goblet cells. Our group has extended this model system by tagging cell populations with fluorescent probes (e.g., green fluorescence protein, or GFP) that allows us to carry out live cell imaging of mixed clonal populations. This model system permits characterization of the ability of more advanced bronchial epithelial cell populations to expand on the growth surface at the expense of less advanced bronchial epithelial cell populations.

Summary of Research Findings

Our long-term plan associated with this Specific Aim was to obtain biopsy specimens from individuals participating in the clinical trial of Project 1 and to compare the differential growth properties in organotypic cultures of bronchial epithelial cells derived from LIFE bronchoscopy positive regions to specimens derived from negative regions. Because of the delay in the initiation of the clinical trial and the limited number of participants in the trial, we initially focused on characterization of the in vitro, organotypic culture growth patterns of a group of established bronchial epithelial cell types, including commercially available normal bronchial epithelial cells; cells from bronchial biopsies that have been immortalized (HCC-BE cells) with transfection of cyclin-dependent kinase 4 (cdk4) and telomerase (h-TERT) by our colleagues, Drs. Minna and Shay; and bronchial epithelial cells derived from Ad12/SV40-transfected bronchial epithelial cells (BEAS 2B) and subsequently immortalized in vivo (1799), transformed with carcinogens (1198), and then selected for a tumorigenic phenotype (1170).

In previous studies, we transfected different cell lines with differing living color probes (e.g., GFP, RFP, and CFP) and characterized their relative ability to grow and expand at the expense of their competitor cell populations. Initial studies used fluorescence microscopy to delineate the spatial domains occupied by the competing cell populations, and laser scanning microscopy was used to determine the manner in which one cell population could expand at the expense of its neighboring population. More recently, we examined the interface dynamics between competing cell populations using live cell imaging and were able to directly observe, in real time, the manner in which one cell population advances at the expense of its neighboring colony.

One problem that we face in carrying out the proposed studies with fresh bronchial cell outgrowths (obtained from biopsies from different lung sites from the same individual and initiated in culture by Dr. Koo’s laboratory) is that the cells will not contain living color markers. Such cells are notoriously difficult to transfect and not amenable to selection using antibiotic treatments. We therefore initiated the development of lentiviral infection technologies that would provide high infection rates and the ability to express genes in slowly growing cell populations. Lentiviral constructs have now been developed that will allow us to label bronchial epithelial cells with GFP, YFP, monoRFP, or CFP in freshly generated cultures. We have acquired a new type of color imaging vector that contains copies of red, cherry, YFP and blue genes on the same vector and that contain lox recombination sequences between the different color sequences. Transfection of cells containing this multicistronic vector with the gene for the cre recombinase leads to recombination and the generation of cells with differing colors (Fig 8).

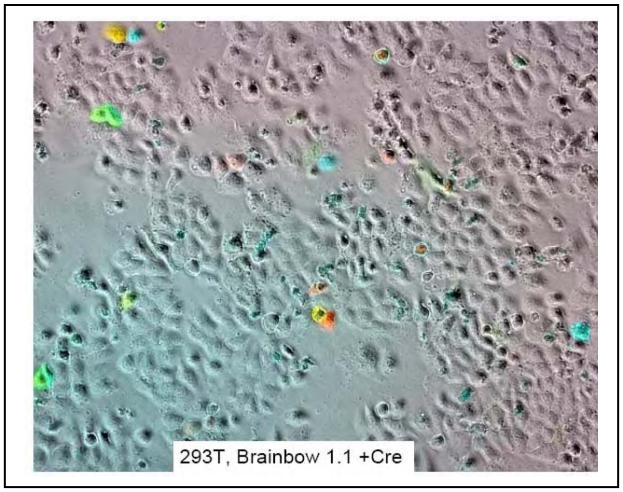


Figure 8. Development of multicolor cell populations using the Brainbow vector 1.1. The cells are first stably transfected with the Brainbow vector 1.1 and then transfected with Cre recombinase to generate heterogeneity of color within the cell population. This approach will permit direct, real-time visualization of preferential subclonal expansion within the cell population.

Outgrowths of the resulting subclones can then be visualized and quantified using our spectral imaging camera, which characterizes the spectral characteristics of each picture element in the image and allows us to distinguish groups of cells differing in color. We are hopeful this new approach will allow us to identify and selectively isolate more aggressive subpopulations that exist within heterogeneous cell populations and characterize the specific genetic and expression changes associated with their more aggressive behavior.

Aim 3 Determine the mechanisms of genetic instability and elucidate the signaling pathways associated with clonal outgrowth of premalignant and malignant bronchial epithelial cells using the organotypic model system.

Our prior studies using chromosome *in situ* hybridization to visualize genetic changes in the bronchial epithelium of current and former smokers suggested that current tobacco exposure was associated with increased levels of ongoing genetic instability (i.e., chromosome polysomy). Upon smoking cessation, while the initiated clonal outgrowths appeared to be maintained over tens of years, the levels of ongoing genetic instability appeared to decrease gradually during the first year following smoking cessation. However, in some cases, we observed evidence for ongoing genetic instability in the bronchial epithelial cells even 10-20 years following smoking cessation. Since nearly half of the newly diagnosed lung cancer cases occur in former smokers, we felt that this finding suggested that an ongoing intrinsic process of genetic instability might exist in the lungs of some former smokers that drives continued genetic evolution toward lung cancer even after cessation of extrinsic carcinogenic exposure.

Our working hypothesis is that years of tobacco exposure induces a chronic damage and wound healing cycle that results both in the accumulation of genetic alterations in the epithelial cells that influences both chromosome stability mechanisms (e.g., loss of cell cycle checkpoint and cell loss mechanisms through loss of p16 expression, p53 mutations, cyclin D1 overexpression, etc) and creates a poor growth environment (e.g., altered stromal signals). The goal of this specific aim was to utilize the lung organotypic model to address this hypothesis *in vitro* utilizing bronchial epithelial cells derived from LIFE bronchoscopically identified “abnormal” and “normal” regions of the lung of current and former smokers participating in the clinical trial of Project 1.

Summary of Research Findings

Previous data suggested that ongoing genetic instability could be detected in the bronchial epithelial cell populations of both current and former smokers leading to the development of subclonal outgrowths. When grown in three-dimensional, lung organotypic cell cultures, we found that bronchial epithelial cells exhibited genetic instability in the form of errors at mitosis, including chromosome breaks, lagging chromosome fragments at anaphase, chromosome bridges, multipolar spindles, and development of bi- and trinucleate cells due to failed completion of cell separation following mitosis (Fig. 9). We noted that the levels of genetic instability increased when cells proliferate away from the basal layer, suggesting dysregulation of cell cycle control when cells try to proliferate in inappropriate spatial regions. We also showed that mitotic instability was associated with increased expression of stress-response proteins, including phosphorylated histone H2AX.

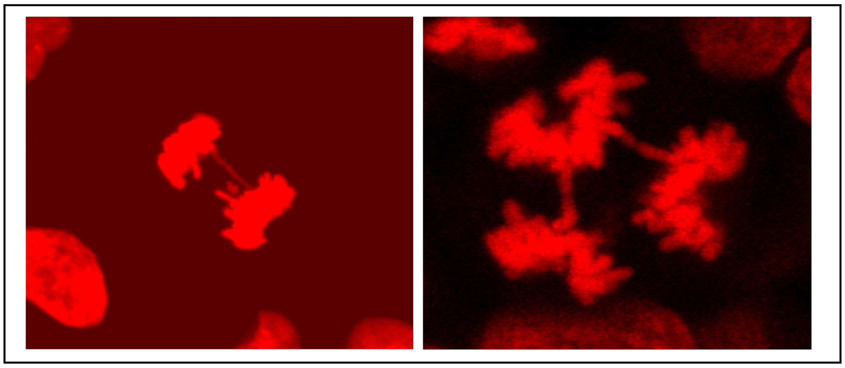


Figure 9. Examples of ongoing genetic instability in bronchial epithelial cells in organotypic cultures. The figure on the left shows an anaphase cell containing a bridge and a lagging fragment. The figure on the right shows a tripolar anaphase. The mitotic figures are stained with propidium iodide, which gives a red-orange fluorescence.

Recent data suggests that genetic instability can occur when cells try to undergo mitosis without proper spatial directionality; the orientation of mitosis may be regulated by its local physical environment. In our organotypic culture system, we noted that normal bronchial epithelial cells preferentially undergo mitosis at the basal layer of the culture and the orientation of the mitotic apparatus is generally parallel to the basal layer, perhaps associated with the attachment to the substratum (Fig. 10). However, we noted that cells further along the multi-step tumorigenesis pathway tended to divide more frequently away from the basement membrane and their mitotic spindle orientation relative to the plane of the basal layer was highly variable (Fig. 11).

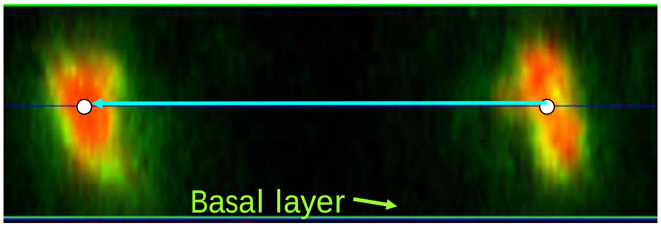


Figure 10. Example of a normal human bronchial epithelial cell in anaphase with its spindle orientation parallel to the plane of the basal layer

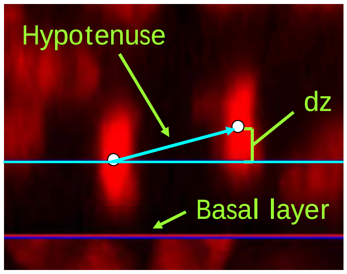


Figure 11. Example of an immortalized human bronchial epithelial cell in anaphase with its spindle orientation at an angle of dz relative to the plane of the basal layer.

To better quantify this abnormal phenomenon, we quantitatively analyzed images of the three dimensional cultures to determine the relationship between changes in mitotic orientation relative to the basal layer plane and the frequencies of mitotic errors (Fig. 12). We found that as the cultures fashioned toward the

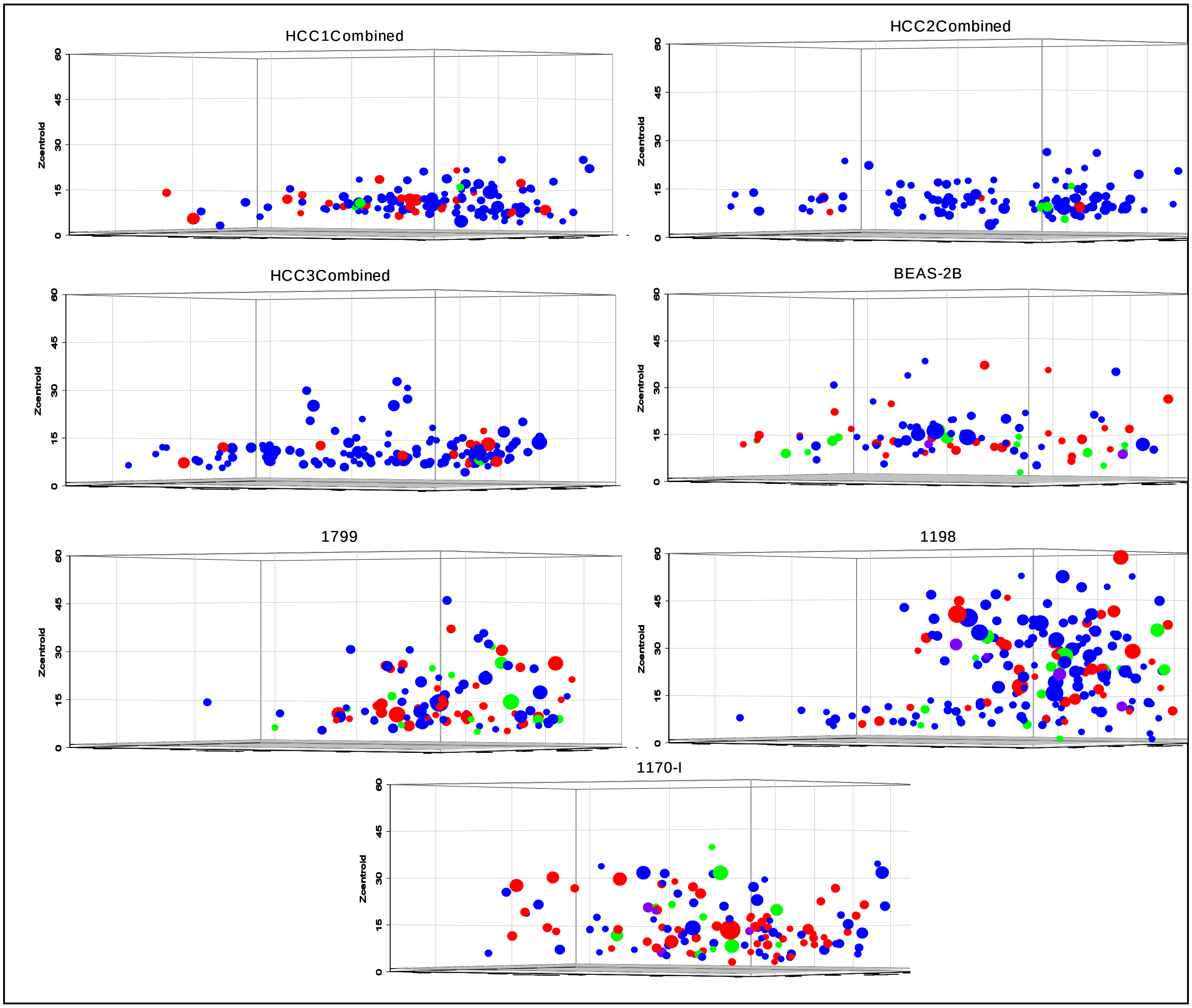


Figure 13. Graphical representation of the relationship between the distance of the anaphase of the anaphase from the basal layer, the orientation of the mitotic spindle relative to the plane of the basal layer, and the presence of mitotic abnormalities in bronchial epithelial cells at various stages of the multistep tumorigenesis process cultured in organotypic cultures. HCC 1, 2 and 3 cells represent normal human bronchial epithelial (NHBE) cells that have been immortalized with cdk4 and hTERT. Beas 2B cells represent NHBE cells that have been immortalized with SV40 large T antigen. 1799 and 1198, cells represent BEAS 2B cells that have been treated with carcinogen, transformed, and further along the multiistep tumorigenesis pathway. 1170I cells are BEAS 2B cells that have been treated with carcinogen and have become tumorigenic. The symbols in the figures are same as those described in the legend of Figure 12.

tumor phenotype, they more frequently exhibited mitoses away from the basal layer and the orientation angle of the mitotic spindle (and degree of mitotic error) increased with distance from the basal layer (Figs. 13 and 14). This result supports the hypothesis that mitotic directionality and fidelity depend on cell-cell and cell-substrate interactions. Cells that undergo mitosis in an inappropriate physical context may improperly orient their mitotic spindles and result in mitotic error and increased levels of genetic instability (Fig 14). We therefore postulate that re-regulating the spatial location of mitotic events in the bronchial epithelium may decrease the rate of ongoing genetic instability and perhaps delay cancer onset in individuals with increased lung cancer risk. This organotypic culture model might be useful in the detection of potential chemopreventive agents that can reregulate the spatial patterns of mitotic events in the bronchial epithelium.

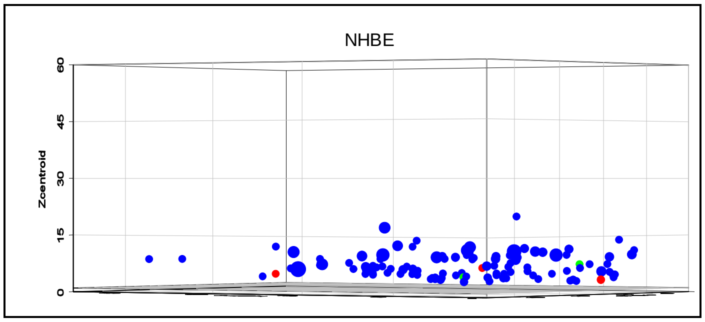


Figure 12. Graphical representation of the relationship between the distance of the anaphase of the anaphase from the basal layer, the orientation of the mitotic spindle relative to the plane of the basal layer, and the presence of mitotic abnormalities in normal human bronchial epithelial cells in an organotypic culture. The Y axis represents the distance of the anaphase from the basal layer. The color of the symbol represents the absence (blue) or presence of an abnormal anaphase, including lagging chromosomes (red), chromosome bridges (green), or multipolar spindles (purple). The size of the symbol represents the angle of the mitotic spindle relative to the plane of the basal layer.

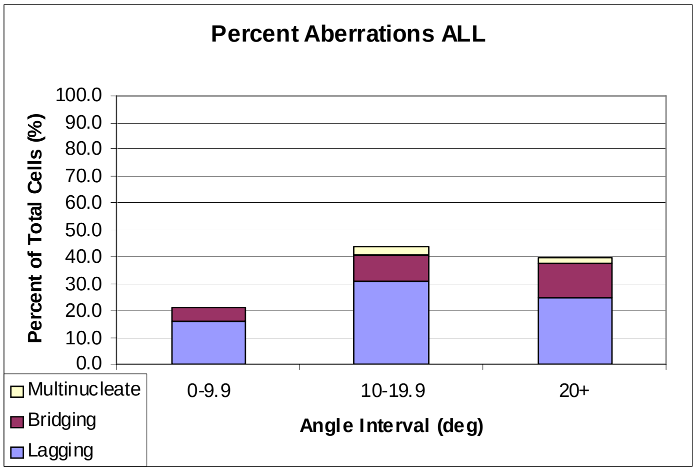


Figure 14. Relationship between deviations in the relative angle of the mitotic spindle relative to the basal layer plane and the frequency of mitotic cells showing anaphase abnormalities.

Aim 4 Characterize the impact of chemopreventive and/or chemotherapeutic agents on early lung tumorigenesis events in reconstructed bronchial epithelium and in the bronchial biopsies of subjects entered onto the clinical trials in Project 1.

The goals of the first three specific aims of this project are essentially to develop and utilize the lung organotypic culture model to identify the factors that control ongoing clonal expansion and genetic instability in the lungs of current and former smokers. The idea behind this fourth specific aim is to integrate the information garnered from the first three specific aims to identify targeted strategies to slow preferential outgrowth of more advanced bronchial epithelial cells and to decrease the levels of ongoing genetic instability. We also proposed to determine whether treatment of these organotypic cultures with the chemopreventve agents used in the clinical trial of Project 1 would slow these aberrant properties *in vitro* and whether results obtained in the organotypic culture model reflected that seen in the lungs of the participants in the clinical trial.

Summary of Research Findings

This aspect of the project is still in development. We have generated lentiviral vectors for inducing living color probes in primary bronchial epithelial cells to examine clonal outgrowths. We have also generated lentiviral vectors for color-marked histone H2B that will permit us to directly monitor mitotic fidelity in real time in organotypic cultures of bronchial epithelial cells, and have developed lentiviral vectors containing either color-marked cyclin D1a or cyclin D1b that will allow us to determine the relative impact of cyclin D1 isoform overexpression on genetic instability. Parallel studies in immortalized oral epithelial cells have demonstrated that cyclin D1 overexpression induces various types of mitotic instability including increased frequencies of chromosome bridges and lagging chromosome fragments and generation of binucleate cells associated with incomplete cell separation at the end of mitosis. Dr. Koo has already established more than 14 pairs of cell strains derived from bronchial biopsies of fluorescence-normal and -abnormal regions of participants in the prevention trial of Project 1. These strains are being stored in the frozen state, and further studies will be undertaken after appropriate lentiviral particles are generated for infection of these cells.

Aim 5. Identify gene expression signatures in epithelial cells detected by LIFE bronchoscopy that determine aggressiveness.

This new aim under the ReVITALization plan will be performed in conjunction with Dr. Wistuba, Core C Director. Our preliminary results from our research in VITAL have shown that epithelial cells isolated from bronchial biopsies of LIFE-abnormal mucosa can be characterized as more aggressive (invasive and migratory) than those of LIFE-normal mucosa. Microarray analysis suggested that several CXCL-chemokine signaling pathways are mainly deregulated in LIFE-abnormal cells. Moreover, we identified that pro-angiogenic ELR+ (glutamic acid, lysine and arginine motif) chemokines were strongly upregulated by inflammatory cytokines in lung cancer cells.

Summary of Research Findings

Inflammatory cytokines are known to play important role in formation of new blood vessels. The proinflammatory cytokine interleukin (IL)-1β has been reported to promote tumor development. We reported that IL-1β up-regulated an array of proangiogenic CXC chemokine genes in the NSCLC cell line A549 and NCI-H1734, as determined by microarray analysis. Conditioned medium from IL-1β–treated A549 and NCI-H1734 cells markedly increased endothelial cell migration. The migration was completely suppressed by neutralizing antibodies against CXCL5 and CXCR2. We also found that IL-1β–induced CXC chemokine gene overexpression in NSCLC cells was abrogated with the knockdown of two transcription factors, CREB or NF-κB. Moreover, the expression of the CXC chemokine genes as well as CREB and NF-κB activity was greatly increased in the tumorigenic NSCLC cell line compared with normal, premalignant immortalized or nontumorigenic cell lines. A disruptor of the interaction between CREB-binding protein and transcription factors such as CREB and NF-κB, 2-naphthol-AS-E-phosphate (KG-501), inhibited IL-1β–induced CXC chemokine gene expression and angiogenic activity in NSCLC. Further studies revealed that KG-501 induced cell cycle arrest and apoptosis of lung cancer cell lines. Based on these findings, we proposed that targeting CREB or NF-κB using small-molecule inhibitors, such as KG-501, holds promise as a preventive and/or therapeutic approach for lung cancer.

Genes in the EGFR-MAPK pathway have been reported as being abnormally regulated in abnormal bronchial epithelial cells obtained from bronchial brush specimens guided by LIFE. During the development of squamous cell carcinoma (SCC) in the lung, bronchial epithelial cells exhibit a progressive series of morphologically distinct changes: hyperplasia, squamous metaplasia, dysplasia, carcinoma in situ, and finally invasive SCC. Here, we investigated molecular mechanisms involved in the initiation of the abnormal differentiation, namely hyperplasia, of bronchial epithelial cells. We demonstrated that ErbB1 ligands, including epidermal growth factor (EGF), TGF-alpha, and amphiregulin, completely disrupted apical-basal polarity and induced hyperplasia of normal human tracheobronchial epithelial (NHTBE) cells. EGF-induced hyperplasia was completely blocked by an EGFR inhibitor, erlotinib, and MEK1/2 inhibitor, U0126, suggesting involvement of MEK-ERK signaling. Further studies showed that EGF substantially upregulated cyclin D1, and that these inhibitors completely blocked the upregulation. Promoter analysis of cyclin D1 revealed that Ap-1 transcription factor regulates the overexpression of cyclin D1. Depletion of AP-1 component c-Jun using siRNA completely abrogated EGF-induced cyclin D1 expression and also inhibited EGF-induced hyperplasia in NHTBE cells. In conclusion, we showed that EGF induced hyperplasia of primary bronchial epithelial cells and AP-1 plays a crucial role in lung carcinogenesis. The results were published in a recent issue of Cancer Prevention Research (October 2008).

Key Research Accomplishments

- Completed processing of 14 pairs of patient samples and obtained the microarray data sets.
- Identified two distinct populations among patient samples, which has helped us to avoid the confounding effect of a mixed sample population and enabled us to find consensus in gene regulation patterns.
- Identified the cytokine-cytokine receptor pathway as one of the major signal pathways that may dictate the differences between the white-light or LIFE abnormal tissue and normal tissue, and could be critical in the early development of lung cancer.
- Located panels of genes that are unanimously regulated in each distinct sample cluster, and found that several of these genes coincide with genes previously reported to be important for tumorigenesis.
- Demonstrated that cells further along the multi-step tumorigenesis pathway tended to divide more frequently away from the basement membrane and their mitotic spindle orientation relative to the plane of the basal layer was highly variable.
- Confirmed that as cultures fashioned toward the tumor phenotype, they more frequently exhibited mitosis away from the basal layer and the orientation angle of the mitotic spindle (and degree of mitotic error) increased with distance from the basal layer.
- KG501, a small molecule inhibitor targeting CREB activity, and neutralizing antibodies against CXCL5 and CXCR2 blocked the migration of vascular endothelial cells induced by inflammatory cytokine.
- Demonstrated that EGFR ligands induce hyperplasia of bronchial cells cultured in an organotypic 3-dimensional culture method.
- Demonstrated that cyclin D1 upregulated by AP1 transcription factor plays a critical role in the hyperplasia of the bronchial epithelial cells and that erlotinib blocks the hyperplasia.

Conclusion

- Cytokine-cytokine receptor pathway was identified as one of the major signal pathways that may dictate the differences between the white-light or LIFE abnormal tissue and normal tissue, and hence could play a critical role in the early development of lung cancer.
- KG501, a small molecule inhibitor targeting CREB activity, and neutralizing antibodies against CXCL5 and CXCR2 hold promise as preventive and/or a therapeutic strategy for lung cancer.
- Demonstrated that re-regulating the spatial location of mitotic events in the bronchial epithelium may decrease the rate of ongoing genetic instability and perhaps delay cancer onset in individuals with increased lung cancer risk.
- EGFR ligands induce bronchial hyperplasia, which can be inhibited by erlotinib.

Project 4: Modulation of Death Receptor-Mediated Apoptosis for Chemoprevention

(Project Leader and co-leader: Shi-Yong Sun, Ph.D.; Fadlo R. Khuri, M.D.)

The objective of Project 4 is to understand the role of death receptor (DR)-mediated apoptotic pathways in lung carcinogenesis, cancer prevention, and therapy in order to develop mechanism-driven combination regimens by modulating DR-mediated apoptosis for chemoprevention and therapy of lung cancer. Following is a summary of our research progress:

Aim 1: To determine whether decoy receptor (DcR) and tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) expression are reduced or lost while DR remains largely expressed and whether procaspase-8 and FLIP expression and Akt activity are increased during lung carcinogenesis.

Summary of Research Findings

c-FLIP is the major inhibitor of death receptor-mediated extrinsic apoptotic pathway. We determined the expression patterns of c-FLIP in a panel of cell lines including immortalized normal, premalignant and malignant (i.e., non-small cell lung carcinoma; NSCLC) bronchial epithelial cell lines. Both long and short forms of c-FLIP were highly expressed in all of the malignant NSCLC cell lines (i.e., 1170-I, H157, A549, H1299, H1792, and HCC827) and the premalignant line 1198 that was exposed to tobacco carcinogens; in contrast, the normal cell lines (HBEC3KT and BEAS-2B) and the premalignant cell line 1799 that was not exposed to tobacco carcinogens expressed very low or undetectable levels of c-FLIP (Fig. 1). Death receptor 5 (DR5), as we speculated, is expressed in both normal and malignant lung cancer cell lines. Bcl-2, a major inhibitor of the mitochondria-mediated intrinsic apoptotic pathway, was also expressed in normal cell lines and the majority of NSCLC cell lines (Fig. 1). Together, these results suggested that c-FLIP may play an important role in regulation of the development of lung cancer and may serve as a therapeutic target for lung cancer.

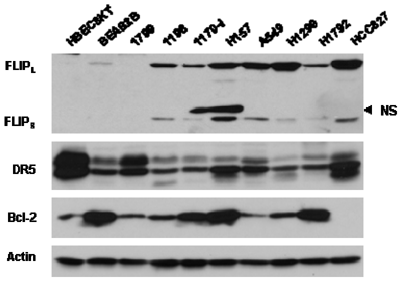


Figure 1. Detection of c-FLIP, DR5 and Bcl-2 expression in normal and transformed bronchial epithelial cell lines. Whole protein lysates were prepared from the indicated cell lines and then subjected to detection of the given protein by Western blot analysis. NS, non-specific spot.

We thus used c-FLIP small interfering RNA (siRNA) to specifically downregulate the expression levels of c-FLIP (both long and short forms) and determined its impact on apoptosis. Knockdown of c-FLIP substantially induced cleavage of caspase-8, caspase-9, caspase-3 and PARP (Fig. 2A), and increased annexin V-positive cells (apoptotic cells) (Fig. 2B). These results clearly indicate that downregulation of c-FLIP alone is sufficient to trigger apoptosis, further supporting that c-FLIP is a promising lung cancer therapeutic target.

We are currently identifying the optimal c-FLIP antibody for IHC so that we can confirm our findings of c-FLIP overexpression in human NSCLC specimens by IHC.

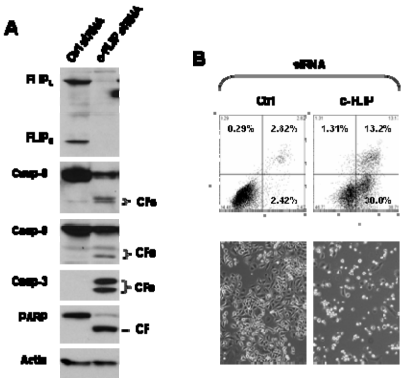


Figure 2. Downregulation of c-FLIP induces apoptosis of lung cancer cells. A549 cells were transfected with control (Ctrl) or c-FLIP siRNA for 48 h and then subjected to preparation of whole cell protein lysates for detection of the indicated proteins by Western blotting (A) or to make single cell suspensions for detection of apoptosis using annexin C staining (B). CF, cleaved fragments.

Aim 2: To establish TRAIL-resistant cell lines from a TRAIL-sensitive lung cancer cell line and determine whether levels of DcRs, DRs, procaspase-8, TRAIL and FLIPs and Akt activity are altered and are associated with cell resistance to TRAIL and DR-inducing agents.

As noted in the previous report, we were not able to demonstrate that the TRAIL-resistant lung cancer cell lines exhibited cross-resistance to some DR-inducing agents. Alternatively, we have focused on addressing the question of whether these agents modulate the DR-mediated apoptotic pathway and, if so, how they modulate the DR-mediated apoptotic pathway and whether the modulations impact apoptosis by these DR-inducing agents.

Summary of Research Findings

A. CHOP-dependent death receptor 5 induction is a major component of SHetA2-induced apoptosis in lung cancer cells.

The flexible heteroarotinoids (Flex-Hets) represent a novel type of atypical retinoids lacking activity in binding to and transactivating retinoid receptors. Preclinical studies have demonstrated that Flex-Hets induce apoptosis of cancer cells while sparing normal cells, and exhibit anticancer activity *in vivo* with improved therapeutic ratios over conventional retinoid receptor agonists. Flex-Hets have been shown to induce apoptosis through activation of the intrinsic apoptotic pathway. The present study has revealed a novel mechanism underlying Flex-Het-induced apoptosis involving induction of death receptor 5 (DR5). The representative Flex-Het SHetA2 effectively inhibited the growth of human lung cancer cells in cell culture and in mice. SHetA2-induced apoptosis, which could be abrogated by silencing caspase-8 expression, indicate that ShetA2 triggers a caspase-8-dependent apoptosis (Fig. 3). Accordingly, SHetA2 upregulated DR5 expression including cell surface levels of DR5 and augmented tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis. Importantly, small interference RNA (siRNA)-mediated blockade of DR5 induction conferred cell resistance to SHetA2-induced apoptosis as well as SHetA2/TRAIL-induced apoptosis (Fig. 4). These results demonstrate that DR5 induction is a key component of apoptosis induced by SHetA2 or by SHetA2 combined with TRAIL. SHetA2 exerted CHOP-dependent transactivation of the DR5 promoter. Consistently, SHetA2 induced CHOP expression, which paralleled DR5 upregulation, whereas siRNA-mediated blockage of CHOP induction prevented DR5 upregulation, indicating CHOP-dependent DR5 upregulation by SHetA2. Collectively, we conclude that CHOP-dependent DR5 upregulation is a key event mediating SHetA2-induced apoptosis. These data have been published in Cancer Res (Lin et al., Cancer Res 2008).

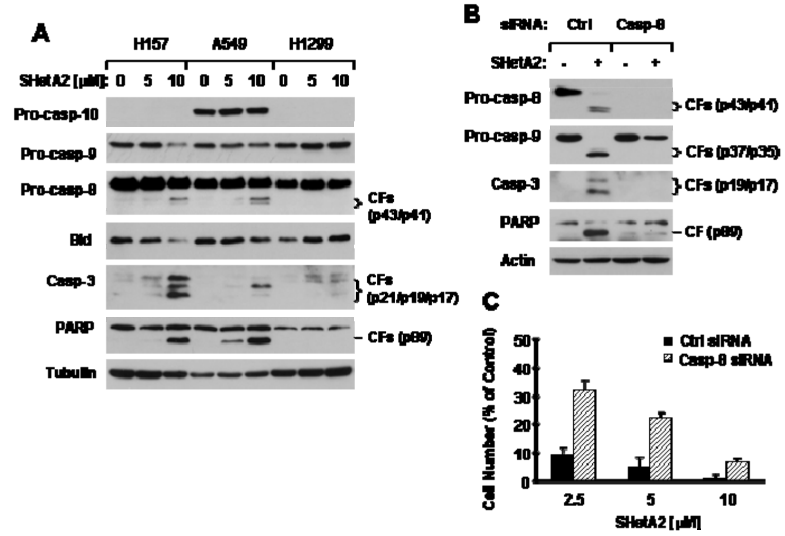


Figure 3. Effects of SHetA2 on cleavage of caspases and their substrates (*A*) and impact of caspase-8 silencing on SHetA2-induced caspase cleavage (*B*) and decrease in cell survival (*C*). *A*, The indicated cell lines were exposed to the given concentrations of SHetA2 for 30 h. The cells were then harvested for preparation of whole-cell protein lysates and subsequent Western blot analysis for detecting cleavage of caspases and their substrates. *B* and *C*, A549 cells were cultured in a 6-well plate and the next day transfected with control (Ctrl) or caspase-8 siRNA. Twenty-four hours after transfection, the cells were re-seeded in a 6-well plate or in a 96-well plate. On the second day, the cells were treated with DMSO or 10 μM SHetA2 (in the 6-well plate) (*B*) or with the indicated concentrations of SHetA2 (in the 96-well plate) (*C*). After 48 h, the cells were subjected to preparation of whole-cell protein lysates and subsequent Western blot analysis (*B*) or to an estimation of cell number using the SRB assay (*C*). Data in *C* are mean of four replicate determinations. Bars, ± SDs. Pro-casp, Pro-caspase; CF, cleaved fragment.

B. Involvement of c-FLIP downregulation in Flex-Het-induced apoptosis and enhancement of TRAIL-initiated apoptosis in lung cancer cells.

The above results demonstrated that upregulation of DR5 plays a critical role in the mechanism of SHetA2-induced apoptosis in human lung cancer cells. The hypothesis of this study was that the mechanism of SHetA2-induced apoptosis requires modulation of additional proteins critical for regulation of apoptosis, including c-FLIP, survivin, XIAP, Bcl-2, Bcl-X_L, Bax and Bim. Western blot analysis demonstrated that c-FLIP and survivin were substantially reduced in all of the tested cell lines exposed to SHetA2 compared to other proteins that were reduced only in a subset of the cell lines tested (Fig. 5). Strikingly, overexpression of c-FLIP, but not survivin, protected cells from SHetA2-induced apoptosis (Fig. 6) and enhancement of TRAIL-initiated apoptosis although knockdown of endogenous survivin did slightly sensitize cells to SHetA2-induced apoptosis. Consistent with these results, small interfering siRNA-mediated reduction of c-FLIP was more effective than survivin downregulation in triggering apoptosis in these cell lines. SHetA2 increased ubiquitination of c-FLIP and the consequent degradation was abrogated by the proteasome inhibitor MG132. Although SHetA2 treatment led to increased c-Jun phosphorylation, the JNK inhibitor SP600125 did not prevent c-FLIP downregulation by SHetA2. Thus, it appears that SHetA2 downregulates c-FLIP levels by facilitating its ubiquitin/proteasome-mediated degradation independent of JNK activation. Collectively, the present study indicates that c-FLIP downregulation is another important component of Flex- Het (SHetA2)-induced apoptosis as well as enhancement of TRAIL-induced apoptosis. This part of the results has been published in Mol Cancer Ther (Lin et al., Mol Cancer Ther 2008).

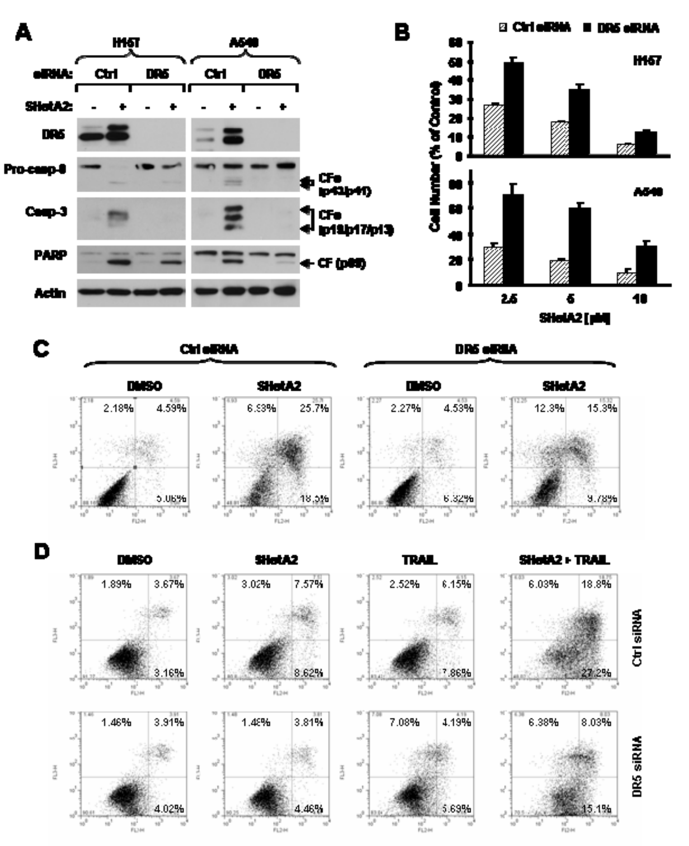


Figure 4. Blockage of DR5 induction (*A*) attenuates SHetA2's ability to activate caspases (A), decrease cell survival (*B*), induce apoptosis (*C*) and augment TRAIL-induced apoptosis (*D*) in human NSCLC cells. *A* and *B*, The indicated cell lines were cultured in 6-well plates and the next day transfected with control (Ctrl) or DR5 siRNA. Twenty-four hours after the transfection, cells were reseeded in 6-well plates (*A*) or 96-well plates (*B*) and treated with 10 μM SHetA2 (*A*) or the indicated concentrations of SHetA2 (*B*). After 48 h (30 h for A549 in *A*), the cells were subjected to preparation of whole-cell lysates and Western blot analysis (*A*) or to the SRB assay for calculation of cell survival (*B*). Data in *B* are mean of four replicate determinations. Bars, ± SDs. Pro-casp, Pro-caspase; CF, cleaved fragment. *C* and *D*, H157 (*C*) or A549 (*D*) cells were cultured in 6-well plates and the next day transfected with control (Ctrl) or DR5 siRNA. Forty-eight hours after the transfection, the cells were exposed to DMSO or 10 μM SHetA2 for 48 h (*C*) or treated with DMSO control, 10 μM SHetA2 alone, 20 ng/ml TRAIL alone or SHetA2 combined with TRAIL for 24 h (*D*). The cells were then harvested for Annexin V assay to detect apoptosis. The percent positive cells in the upper right and lower right quadrants were added to yield the total of apoptotic cells.

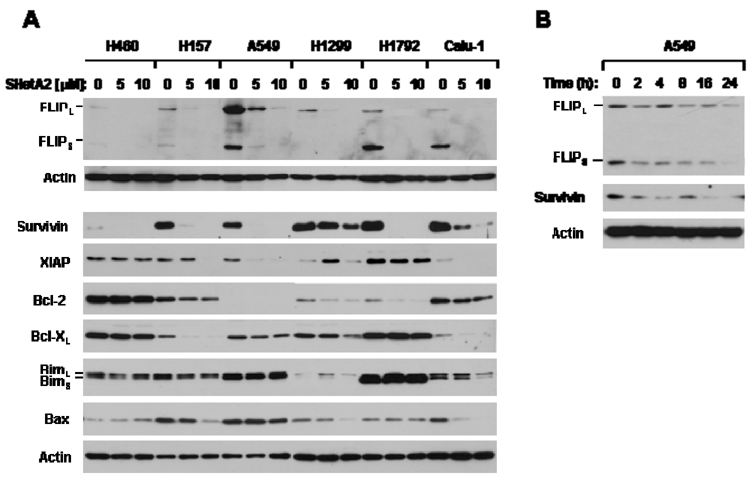


Figure 5. Modulatory effects of SHetA2 on c-FLIP, survivin, XIAP, Bcl-2, Bcl-X_L, Bim and Bax in human NSCLC cell lines. The indicated cell lines were treated with the indicated concentrations of SHetA2 for 16 h (*A*) or with 5 μM SHetA2 for the given time as indicated (*B*). The cells were then subjected to preparation of whole-cell protein lysates and subsequent Western blot analysis.

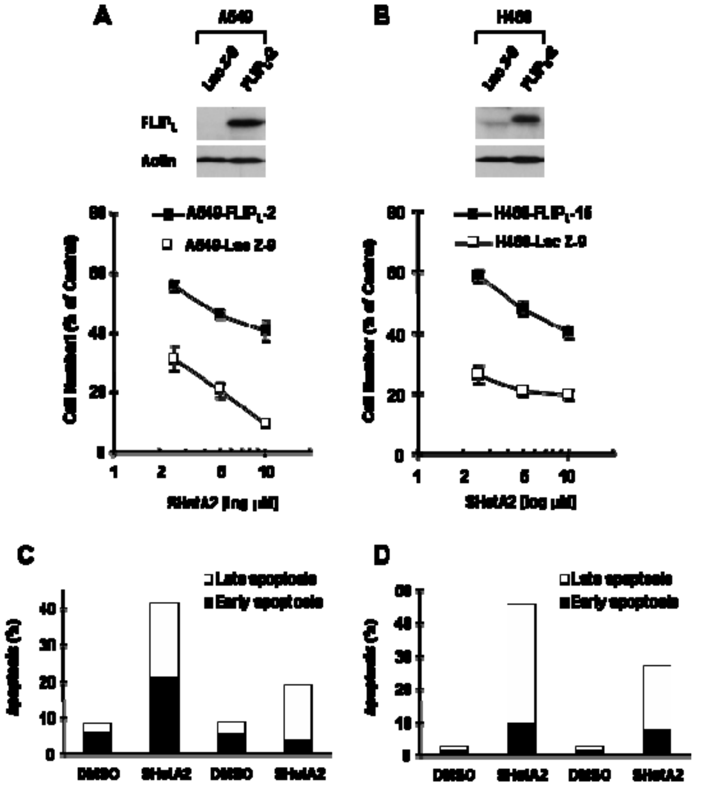


Figure 6. Enforced expression of ectopic c-FLIP protects cells from SHetA2-induced cell number decrease (*A* and *B*) and apoptosis (*C* and *D*). *A* and *B*, The indicated transfectants were seeded in 96-well plates and treated with the indicated concentrations of SHetA2 ranging from 2.5 to 10 μM. After 48 h, the cells were subjected to the SRB assay for measurement of cell number. Data are the means of four replicate determinations. Bars; ± SDs. *C* and *D*, The indicated transfectants were treated with 10 μM SHetA2 for 48 h and then harvested for detection of apoptotic cells using Annexin V staining. The percent positive cells in the upper right and lower right quadrants were added to yield the total of apoptotic cells.

C. Coupling of endoplasmic reticulum (ER) stress to CDDO-Me-induced up-regulation of death receptor 5 via a CHOP-dependent mechanism involving JNK activation.

The synthetic triterpenoid methyl 2-cyano-3,12-dioxoolean-1,9-dien-28-oate (CDDO-Me) is in Phase I clinical trials as a novel cancer therapeutic agent. We previously demonstrated that CDDO-Me induces c-Jun N-terminal kinase (JNK)-dependent death receptor 5 (DR5) expression and augments death receptor-induced apoptosis. The current study focused on addressing how CDDO-Me induces JNK-dependent DR5 expression. Analysis of DR5 promoter

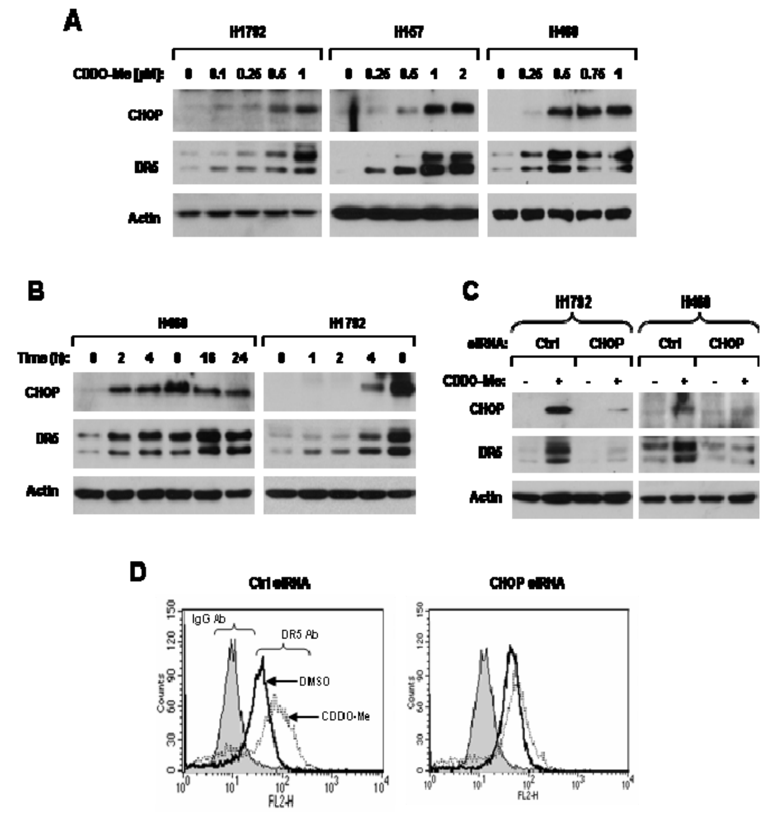


Figure 7. CDDO-Me induces CHOP-dependent DR5 expression. A) The CDDO-Me induces dose-dependent DR5 expression. The indicated cells lines were treated with the indicated concentrations of CDDO-Me for 12 h. **B**) CDDO-Me induces time-dependent DR5 expression. H460 and H1792 cells were treated with 1 μM CDDO-Me for the times as indicated. **C**) Blockade of CHOP induction abrogates CDDO-Me-induced DR5 upregulation. Both H1792 and H460 cells lines were transfected with control (Ctrl) or CHOP siRNA. After 48 h, the cells were treated with 1 μM CDDO-Me for 12 h (H1792) or with 0.5 μM CDDO-Me for 8 h (H460). After the aforementioned treatments (**A**-**C**), the cells were subjected to preparation of whole-cell protein lysates and subsequent Western blot analysis. **D**) Blockade of CHOP induction inhibits CDDO-Me-induced increase in cell surface DR5. H1792 cells were transfected with control (Ctrl) or CHOP siRNA. After 48 h, the cells were treated with 1 μM CDDO-Me for 12 h and then subjected to staining of cell surface DR5 and subsequent flow cytometry. Ab, antibody.

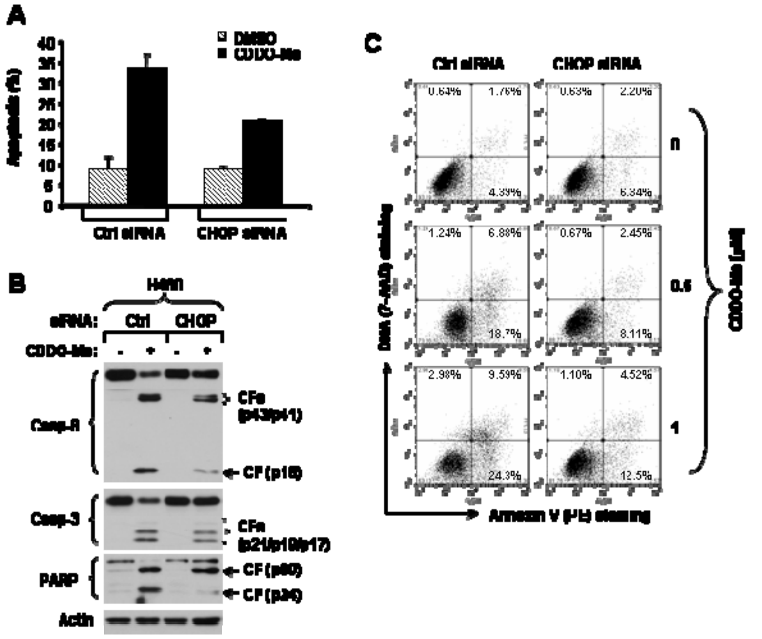


Figure 8. Blockade of CHOP induction diminishes CDDO-Me's ability to induce apoptosis. A) H1792 cells were plated in 6-well cell culture plates and transfected with control (Ctrl) or CHOP siRNA. After 48 h, the cells were treated with the 1 μM CDDO-Me for 12 h and then subjected to flow cytometric analysis for sub-G1 populations. Columns, means of duplicate assays; Bars, ± SE. **B** and **C**) H460 cells were plated in 6-well cell culture plates and transfected with control (Ctrl) or CHOP siRNA. After 48 h, the cells were treated with the 0.5 μM CDDO-Me for 8 h and then subjected to preparation of whole-cell protein lysates and Western blot analysis (**B**). In addition, the cells were also harvested for annexin V staining and flow cytometric analysis of apoptotic cells (**C**). The percent positive cells in the upper right and lower right quadrants were added to yield the total of apoptotic cells.

regions defines that the CHOP binding site is responsible for CDDO-Me-induced transactivation of the DR5 gene. Consistently, CDDO-Me induced DR5 expression and parallel CHOP upregulation. Blockade of CHOP upregulation also abrogated CDDO-Me-induced DR5 expression (Fig. 7). These results indicate that CDDO-Me induces CHOP-dependent DR5 upregulation. Moreover, the JNK inhibitor SP600125 abrogated CHOP induction by CDDO-Me, suggesting a JNK-dependent CHOP upregulation by CDDO-Me as well. Importantly, knockdown of CHOP attenuated CDDO-Me-induced apoptosis, demonstrating that CHOP induction is involved in CDDO-Me-induced apoptosis (Fig. 8). Additionally, CDDO-Me increased the levels of Bip, phosphorylated eIF2α, IRE1α, and ATF4, all of which are featured changes during endoplasmic reticulum (ER) stress. Furthermore, salubrinal, an inhibitor of ER stress-induced apoptosis, inhibited JNK activation and upregulation of CHOP and DR5 by CDDO-Me and protected cells from CDDO-Me-induced apoptosis (Fig. 9). Thus, ER stress appears to be important for CDDO-Me-induced JNK activation, CHOP and DR5 upregulation, and apoptosis. Collectively, we conclude that CDDO-Me triggers ER stress, leading to JNK-dependent, CHOP-mediated DR5 upregulation and apoptosis. These data have been published in Cancer Res (Zou et al., Cancer Res 2008).

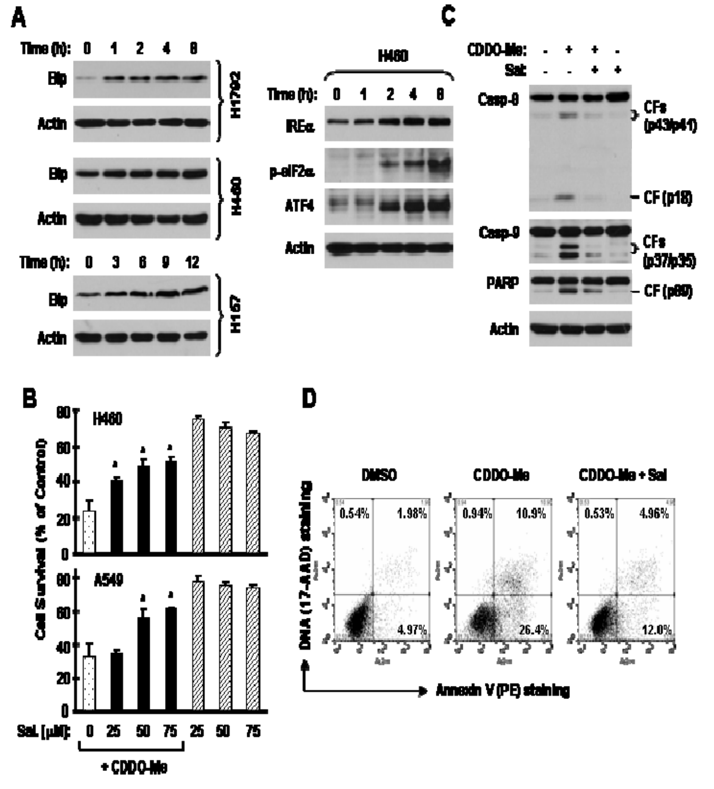


Figure 9. CDDO-Me increases the levels of ER stress marker proteins (A) and induces ER stress-mediated apoptosis (B-D). **A**) The indicated cell lines were treated with 1 μM CDDO-Me. After the indicated times, the cells were harvested and subjected to preparation of whole-cell protein lysates for detection of the indicated proteins by Western blot analysis. **B**) The indicated cell lines were seeded in 96-well cell culture plates and treated the next day with DMSO, 1 μM CDDO-Me alone, the given concentrations of salubrinal (Sal) alone, and the respective combination of CDDO-Me with each of the given concentrations of salubrinal. After 24 h, the cells were subjected to the SRB assay for estimation of cell numbers. Columns, means of four replicate determinations; Bars, ± SDs; * *P* < 0.001 compared with CDDO-Me alone by one-way ANOVA analysis using GraphPad InStat 3 software (GraphPad Software, San Diego, CA). **C** and **D**) H460 cells were seeded in 6-well plates and treated the next day with DMSO, 1 μM CDDO-Me alone, 75 μM salubrinal (Sal) alone, and their combination. After 6 h, the cells were harvested for preparation of whole-cell protein lysates for detection of the indicated proteins by Western blot analysis (**C**) or for detecting annexin V-positive cells using flow cytometry (**D**). CFs, cleaved forms.

Aim 3: To determine whether suppression of PI3K/Akt activity sensitizes premalignant and/or malignant airway epithelial cells to apoptosis induced by DR-induced agents via enhancement of TRAIL/DR-mediated mechanism.

Summary of Research Findings

A. Perifosine inhibits Akt/mTOR signaling through facilitating degradation of major components in mTOR pathway and induces autophagy.

The Akt inhibitor, perifosine, is an alkylphospholipid exhibiting antitumor activity as demonstrated in both preclinical studies and clinical trials. This activity is partly associated with its ability to inhibit Akt activity. It has been shown that mTOR axis plays a critical role in regulation of cell proliferation and survival, primarily through functioning both downstream and upstream of Akt. The current study reveals a novel mechanism by which perifosine inhibits Akt and mTOR axis. In addition to inhibition of Akt, perifosine inhibited the assembly of both mTOR/raptor and mTOR/rictor complexes. Strikingly, perifosine reduced the levels of Akt and other major components including mTOR, raptor, rictor, p70S6K, S6, 4E-BP-1 and Rheb1 in the mTOR axis through promoting their degradation because the proteasome inhibitor MG132 prevented the reduction of these proteins by perifosine. These results suggest that perifosine inhibits the mTOR axis through degradation of major components in this axis, showing a different mechanism from inhibition of mTOR signaling by classical mTOR inhibitors such as rapamycin. Moreover, perifosine substantially increased the levels of type II LC3, a hallmark of autophagy, in addition to increasing PARP cleavage, suggesting that perifosine induces both apoptosis and autophagy (Fig. 10). Collectively, we conclude that perifosine inhibits mTOR signaling and induces autophagy, highlighting a novel mechanism accounting for perifosine's anticancer activity.

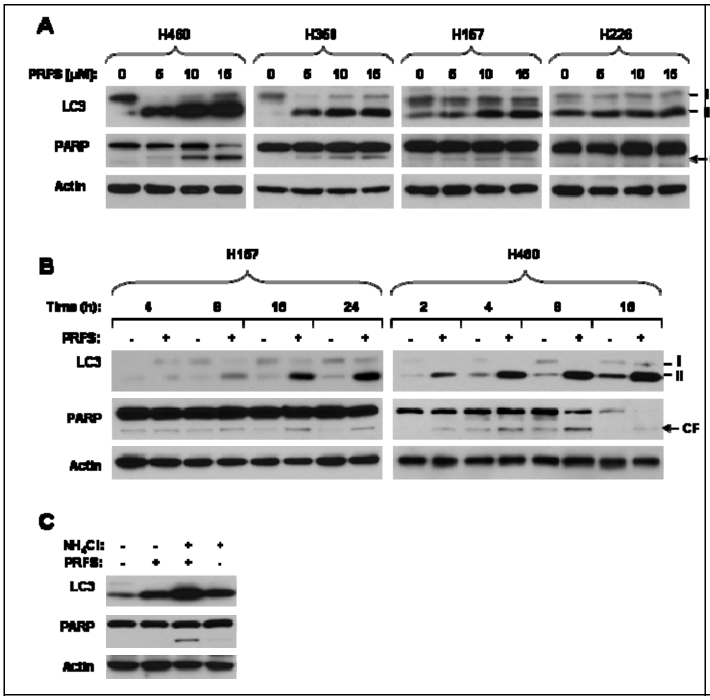

A
H460
H358
H157
H226
PRFS (μM): 0 5 10 15 0 5 10 15 0 5 10 15 0 5 10 15
LC3
PARP
Actin
I
II
CF
B
H157
H460
Time (h): 4 8 16 24 2 4 8 16
PRFS: - + - + - + - + - + - + - + - +
LC3
PARP
Actin
I
II
CF
C
NH4Cl: - - + +
PRFS: - + + -
LC3
PARP
Actin

Figure 10. Perifosine increased the levels of both LC3-II and cleaved PARP (A and B), which are further increased when co-treated with perifosine and NH_4Cl (C). *A* and *B*, The indicated cell lines were treated with the given concentrations of perifosine (PRFS) for 8 h (*A*) or with 10 μM (H157) or 5 μM (H460) perifosine for the given times (*B*). *C*, H460 cells were pre-treated with 20 mM NH_4Cl for 30 min and then co-treated with 5 μM perifosine for additional 4 h. After the aforementioned treatment, the cells were then harvested for preparation of whole cell protein lysates and subsequent Western blot analysis for detection of the indicated proteins.

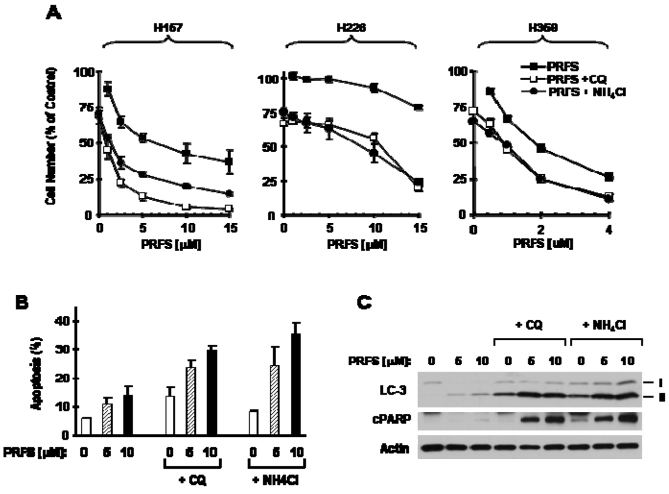


Figure 11. Inhibition of autophage with lysosomal inhibitors enhances perifosine's apoptosis-inducing activity. *A*. The indicated cell lines were seeded on 96-well plates and on the second day treated with the given concentrations of perifosine (PRFS) alone, CQ or NH4Cl alone, or the combination of CQ with PRFS or NH4Cl with PRFS. After 48 h, the cells were subjected to the SRB assay for estimating cell numbers. *B* and *C*, H157 cells were seeded on 10 cm diameter cell culture dishes and treated with on the second day with PRFS alone (5 or 10 uM), CQ or NH4Cl alone, or the combination of PRFS with CQ or NH4Cl. After 48h, the cells were harvested for annexin V staining of apoptotic cells (*B*) or for Western blot analysis of the indicated proteins (*C*).

It has been suggested that autophagy is a survival mechanism that protects cells from being killed though apoptosis. Thus, we further investigated the impact of autophagy on perifosine's apoptosis-inducing activity. We found that inhibition of autophagy with various lysosomal inhibitors enhanced perifosine-induced apoptosis (Fig. 11), suggesting that autophagy induced by perifosine is a survival mechanism, and that its inhibition will enhance perifosine's apoptosis-inducing activity. We are currently studying whether inhibition of autophagy enhances perifosine's anticancer activity in animal xenograft models.

Aim 4: To determine whether DRs, DcRs, c-FLIP, and procaspase-8 serve as biomarkers for lung cancer chemoprevention and therapy.

Summary of Research Findings

This work will be initiated upon receipt of sufficient tissue slides from the clinical trial.

Key Research Accomplishments

1. c-FLIP is overexpressed primarily in transformed malignant lung epithelial cell lines and may serve as a promising therapeutic target for prevention and therapy of lung cancer.
2. The synthetic atypical retinoid SHetA2 induces apoptosis and enhances TRAIL-induced apoptosis involving up-regulation of DR5 and down-regulation of c-FLIP.
3. CDDO-Me-induced up-regulation of death receptor 5 is coupled with endoplasmic reticulum (ER) stress via a CHOP-dependent mechanism involving JNK activation.
4. Perifosine inhibits Akt/mTOR signaling and subsequently induces autophagy. Inhibition of this autophagy augments perifosine's apoptosis-inducing activity, thus suggesting a therapeutic strategy of combining perifosine with an autophagy inhibitor.

Conclusion

Appropriate modulation of the extrinsic death receptor-mediated apoptotic pathway such as upregulation of DR5 and/or reduction of c-FLIP levels by small molecules may eliminate premalignant or malignant lung epithelial cells via promoting apoptotic cell death to achieve cancer chemopreventive and therapeutic goals.

The potential of the modulation of DR5 and c-FLIP as predictive biomarkers for certain drugs in the clinic warrants further investigation. Inhibition of autophagy may be a good strategy to augment certain anticancer agents' anticancer activity.

Project 5: Molecular Strategies Targeting the AKT Signaling Pathway for Lung Cancer Chemoprevention and Therapy

(PI and co-PI: Ho-Young Lee, Ph.D., Edward S. Kim, M.D.)

Our goal is to find novel chemopreventive/therapeutic agents that can prevent lung carcinogenesis effectively. Results from our work and others' have demonstrated that Akt, which has a clear role in cellular survival and transformation, is constitutively active in premalignant and malignant HBECs and in NSCLC cell lines. These findings suggest an importance of PI3K/Akt signaling pathway in lung carcinogenesis. The purpose of our studies is to determine whether activation of Akt induces malignant transformation of HBE cells and to develop novel agents inhibiting Akt activity as a strategy to prevent lung carcinogenesis.

Aim 1 Develop a retroviral vector expressing constitutively active Akt and characterize the *in vitro* and *in vivo* effects of Akt activation on the malignant transformation of HBE cells.

Summary of Research Findings

As we reported last year, we finished constructing retroviral vectors expressing constitutively active or dominant negative Akt1, 2, or 3. The Akt constructs were confirmed by sequencing and Western blot analysis. Viral titers have been determined by the colony formation analysis. To analyze roles of Akt 1, 2 or 3 in the survival of HBE cells, BEAS2B cells were stably transfected with retroviral vectors expressing constitutively active Akt 1, 2, 3. We have been analyzing the survival of the cells in the presence of H2O2, and found that Akt1, 2, and 3 protected HBE cells from cell death (data not shown).

Aim 2 Evaluate the ability of chemopreventive agents used in VITAL trials (gefitinib, erlotinib, SCH66336, and celecoxib, alone and in combination) to inhibit Akt activity and induce apoptosis in transformed HBE and NSCLC cell lines.

Summary of Key Research Findings

Due to increased expression of IGFs detected in human preneoplastic bronchial tissues, in which mutations of *p53* or *KRAS* frequently occur (1), we assessed whether genetic alterations of *p53* or *KRAS* (V12) were associated with increased expression of IGFs and IGF-1R activation. We analyzed immortalized HBE cells (HBEC3) and derivatives, p53 short-interfering RNA-(siRBA)-expressing cells (HBEC3/p53i), or RAS^{v12}-expressing (HBEC3/RAS^{v12}) HBEC3 cells, which harbor several tumorigenic characteristics but are not yet fully malignant. Thus, these cells mimic the premalignant stage human lung carcinogenesis (2). Notably, both HBEC3/RAS^{v12} and HBEC3/p53i cells showed markedly increased IGF-1 and IGF-2 mRNA expression (Figure 1A) and IGF-1R phosphorylation at Tyr1131 (Fig. 1B) and Tyr1135/1136. These cells also showed markedly increased Akt phosphorylation (Ser 473) compared to control HBEC3 cells.

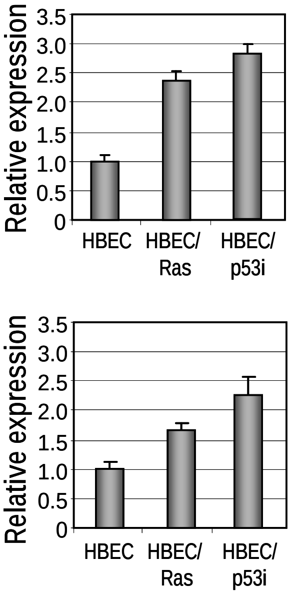


Figure 1A. Expression of IGF-1 and IGF-2 from HBEC3, HBEC3/RASv12 and HBEC3/p53i cells cultured without supplement using quantitative real-time polymerase chain reaction

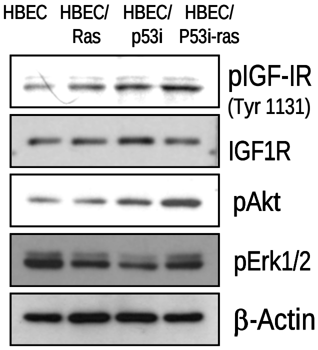


Figure 1B. Enhanced IGF signaling in the HBEC3, HBEC3/RASv12, HBEC3/p53i and HBEC3/ RASv12p53i cells is shown with Western blotting of the markers of the signaling.

Despite the fact that TCs are genotoxic and may induce cell death, HBE cells have shown increased viability when exposed to NNK (3). Thus, we sought to determine whether activated IGF-1R exerted protective effects on TC-exposed lung epithelial cells by analyzing the viability of NNK-exposed HBEC3 cell derivatives. HBEC3 cell derivatives showed markedly increased viability when exposed to NNK in the absence of EGF (Fig. 2A), with the greatest increase seen in p53 RNAi– and RAS^{v12}-expressing HBEC3 cells. Moreover, p53 RNAi– and RAS^{v12}-expressing HBEC3 cells showed the greatest increase in the numbers of foci formed upon reaching a confluent state (Fig. 2B) and anchorage-independent colony forming (Fig. 2C). The NNK-induced increases in viability, foci forming ability and anchorage-independent colony forming activity induced by exposure in p53 RNAi– and RAS^{v12}-expressing HBEC3 cells were almost completely suppressed by PQIP (an IGF-1R TKI). These findings indicated the importance of autocrine IGFs in TC-induced HBE cell transformation.

Dependence of TC-exposed HBE cells IGF-1R signaling for maintaining transformed phenotypes

The findings above showed an important role for IGF-1R–mediated signaling in the transformation of HBE cells, especially that stimulated by TCs. We therefore questioned whether TC-exposed premalignant and malignant HBE cells are dependent on IGF-1R signaling for sustained proliferation and survival. We evaluated the effects of the inhibition of IGF-1R signaling on an *in vitro* model of progressive lung carcinogenesis. The model was composed of the following cell lines: 1799 (a derivative of SV40 large tumor antigen-immortalized bronchial epithelial BEAS2B exposed to beeswax), used as a control; non-tumorigenic 1198 (a derivative of BEAS2B exposed to cigarette-smoke condensate); and tumorigenic 1170-I (a derivative of BEAS2B exposed to cigarette-smoke condensate) (4, 5). The 1198 and 1170-I cells exhibited greater levels of IGF-1R and IRS-1 phosphorylation (Fig. 3A) than 1799 cells. The 1198 and 1170-I cells infected with adenovirus-expressing dominant-negative IGF-1R (Ad-dnIGF-1R) showed a dose-dependent decrease in cell viability (Fig. 3B) and anchorage-dependent colony formation (Fig. 3C) compared to those infected with a control virus (Ad-EV). In contrast, 1799 cells were not affected by the Ad-dnIGF-1R treatment. 1198 and 1170-I cells (Fig. 3D) showed significantly decreased viability in response to the IGF-1R TKI (PQIP), but not to EGFR TKI (erlotinib) when treated in the absence of a growth factor. Neither PQIP nor erlotinib affected the viability of 1799 cells. These findings indicated that IGF -1R activation is required for the transformation of TC-exposed HBE cells.

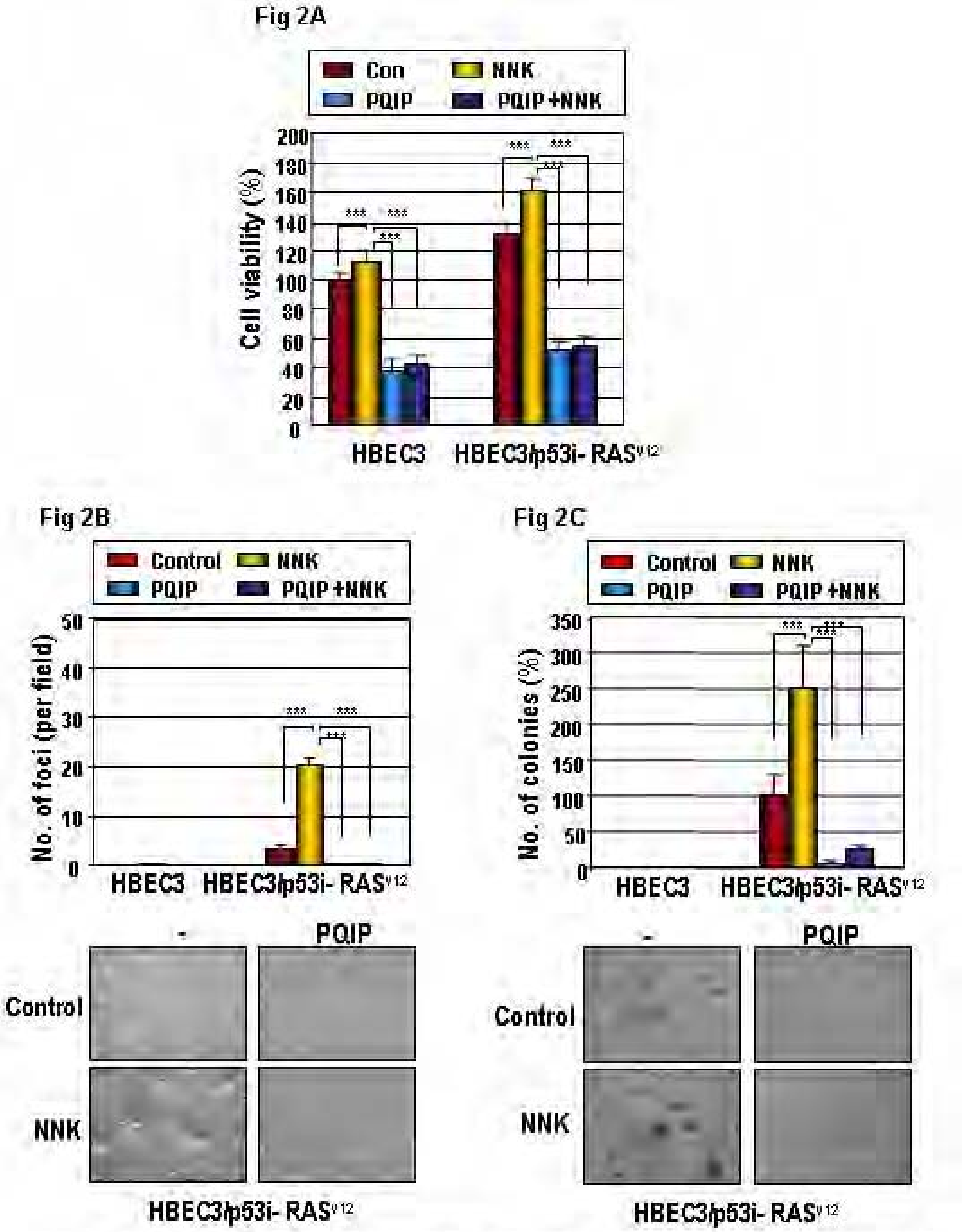


Figure 2. Insulin-like growth factors (IGFs) and tobacco carcinogens synergizing to stimulate the transformation of human bronchial epithelium (HBE) cells. HBEC3 cells and a variant expressing oncogenic K RASv12 (Ras) and knockdown of p53 by RNAi (p53i) were untreated or treated with NNK (1 μM, 5 μM) as indicated and then continuously treated with PQIP (1 μM) or no treatment for 3 days in the absence or presence of NNK. MTT (A, B) and foci-forming assays (C) were performed. A representative picture of foci formation of HBEC3 cells expressing K-RASv12 and p53 RNAi is shown. Data are means ± standard error of the mean. Statistical significance of differences was determined using a two-tailed Student's t test. * ($P < 0.05$); **$P < 0.01$; ***$P < 0.001$.

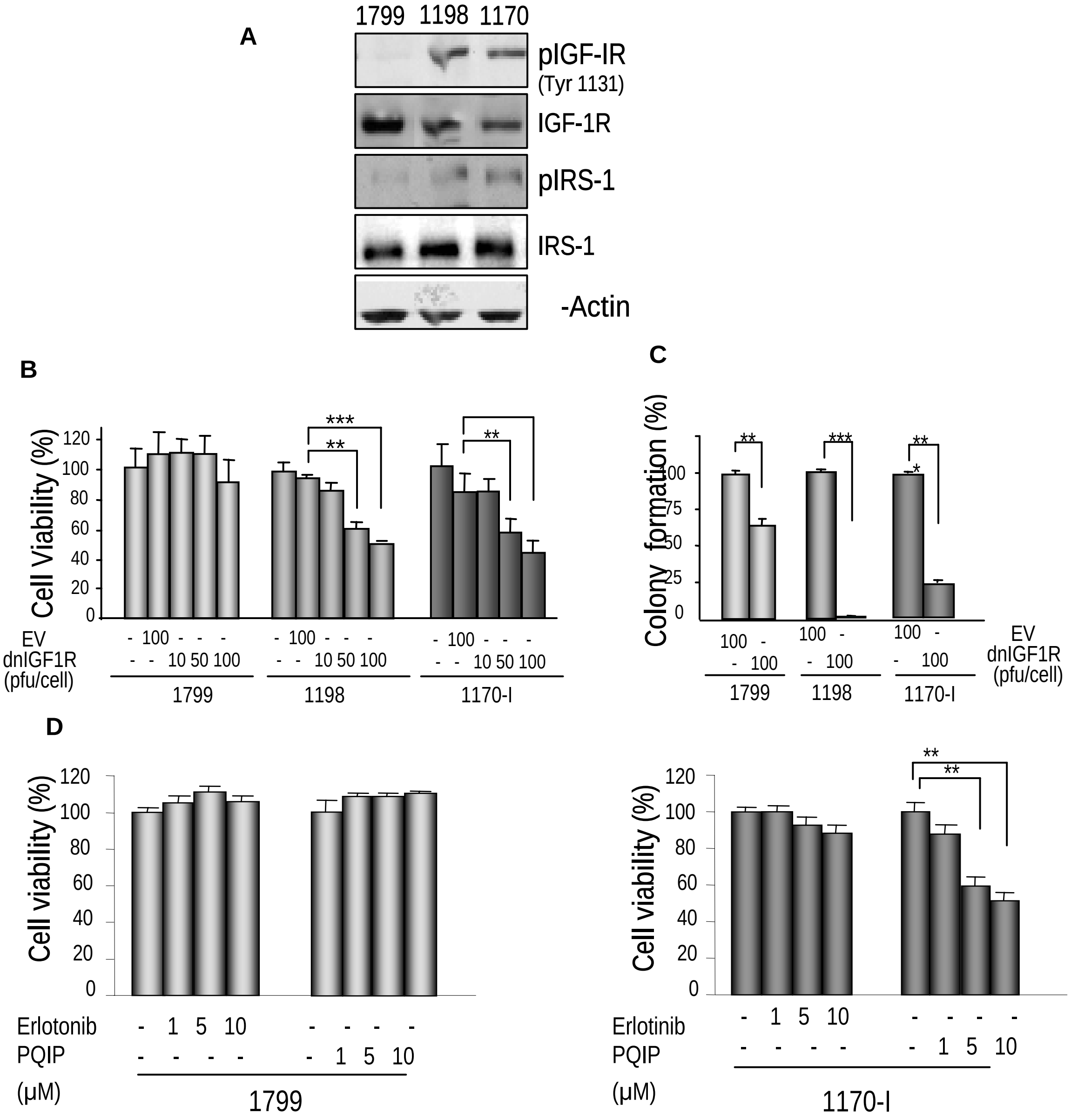


Figure 3. Effects of insulin-like growth factor-1 receptor (IGF-1R) signaling blocking on the survival of tobacco carcinogens (TC)-induced in vitro lung carcinogenesis model cell lines. A) Western blotting of the IGFR signaling molecules in 1788, 1198, and 1170-I cells. Compared to immortalized 1799 cells, the 1198 and 1170-I cells showed highly activated p-IGFR and p-IRS. B, C) Effects of IGF-1R blocking with dnIGF-1R adenovirus on the survival (B) and colony formation (C) of 1799, 1198, and 1170-1 cells. The cells were uninfected or infected with the indicated dose of empty adenovirus or adenovirus expressing dominant negative IGF-1R (dnIGF-1R). After 3 days of incubation in the absence of a growth factor, cells were subjected to the MTT assay (B) and anchorage-dependent colony formation analysis (C). The two transformed cells depend on the IGFR signaling for the survival and colony formation; the normal-like, immortalized cells 1799 do not require IGFR signaling. D) Effects of IGF-1R blocking with PQIP on the survival of 1799 and 1170-I cells. The cells were treated with erlotinib or PQIP at the indicated concentration and subjected to an MTT assay after 3 days. Data are means ± standard deviation of at least three independent experiments. Statistical significance of differences was determined using a two-tailed Student's t test. Differences between groups were considered significant when * P < 0.05; **P < 0.01; and *** P < 0.001.

Aim 3 Determine whether Akt is activated in bronchial specimens from enrolled patients in VITAL trials and whether treatment with chemopreventive agents suppresses Akt level or activity in these patients.

Summary of Research Findings

Our study assessed the role of IGF-1R signaling in lung carcinogenesis. We measured the expression levels of IGF-1, IGF-2, IGF-1R and phospho-IGF-1R in tissue microarrays comprising 367 biopsy specimens of normal, hyperplasic, squamous metaplastic (SQM), low-grade dysplacsia (L-DYS), and high-grade dysplasia (H-DYS) bronchial tissue specimens (Supplementary Table 1 shows patient demographics). IGF-1, IGF-2, and IGF-1R expressions were primarily cytoplasmic and less frequently nuclear (Fig. 1A). We observed significantly higher expression of IGF-1 ($P < 0.0001$) and IGF-2 ($P = 0.004$) in H-DYS than in normal bronchial specimens (Figs. 4A and 4B). IGF-1R is primarily activated by its cognate ligands (IGF-1 and IGF-2) with 2 to 15 times less affinity for IGF-2 than for IGF-1. IGFs also bind to the insulin receptor (IR) with roughly 100 times less affinity than insulin (6). We assessed whether increased levels of the IGFs were associated with activation of IGF-1R signaling by performing immunohistochemical analysis using an antibody that detects phosphorylated IGF-1R (pIGF-1R/IR) (Tyr 1162/1163); the staining appeared in the cell membranes, cytoplasm, and nuclei. Staining in the membrane was significantly higher in H-DYS than in normal, hyperplastic, SQM, and L-DYS bronchial specimens (Fig. 4B) and correlated well with the levels of IGF-1 and IGF-2 (Fig. 4C).

We assessed the role of survivin, one of the inhibitors of apoptosis protein regulated by IGF-1R/Akt signaling pathway, in lung carcinogenesis. We investigated the effects of the tobacco component nicotine and its related carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) on survivin expression in normal human bronchial epithelial (HBE) cells and examined the role of survivin in the malignant transformation of HBE cells induced by these components. We found that survivin mRNA expression was detected in 41% (7 of 17) of bronchial brush specimens from heavy smokers (at least 20 pack/year). Nicotine and NNK increased survivin mRNA and protein expression levels in primary cultured normal HBE cells and immortalized HBE cells. Nicotine and NNK stimulated the Akt/mammalian target of rapamycin (mTOR) pathway in normal HBE cells, leading to increased *de novo* synthesis of survivin protein. Induced survivin expression increased the survival potential of the cells, which was blocked by transfection with survivin-specific small interfering RNA. Small interfering RNA-induced down-regulation of survivin expression also suppressed the tumorigenic potential of premalignant and malignant HBE cells exposed to the tobacco components. These findings suggest that NNK and nicotine induce survivin protein synthesis in normal HBE cells by activating the Akt/mTOR pathway and, thus, blockade of the pathway effectively inhibits the tobacco-induced malignant transformation of HBE cells. These results are published in Jin et al., Carcinogenesis 29(8):1514-1622 (2008).

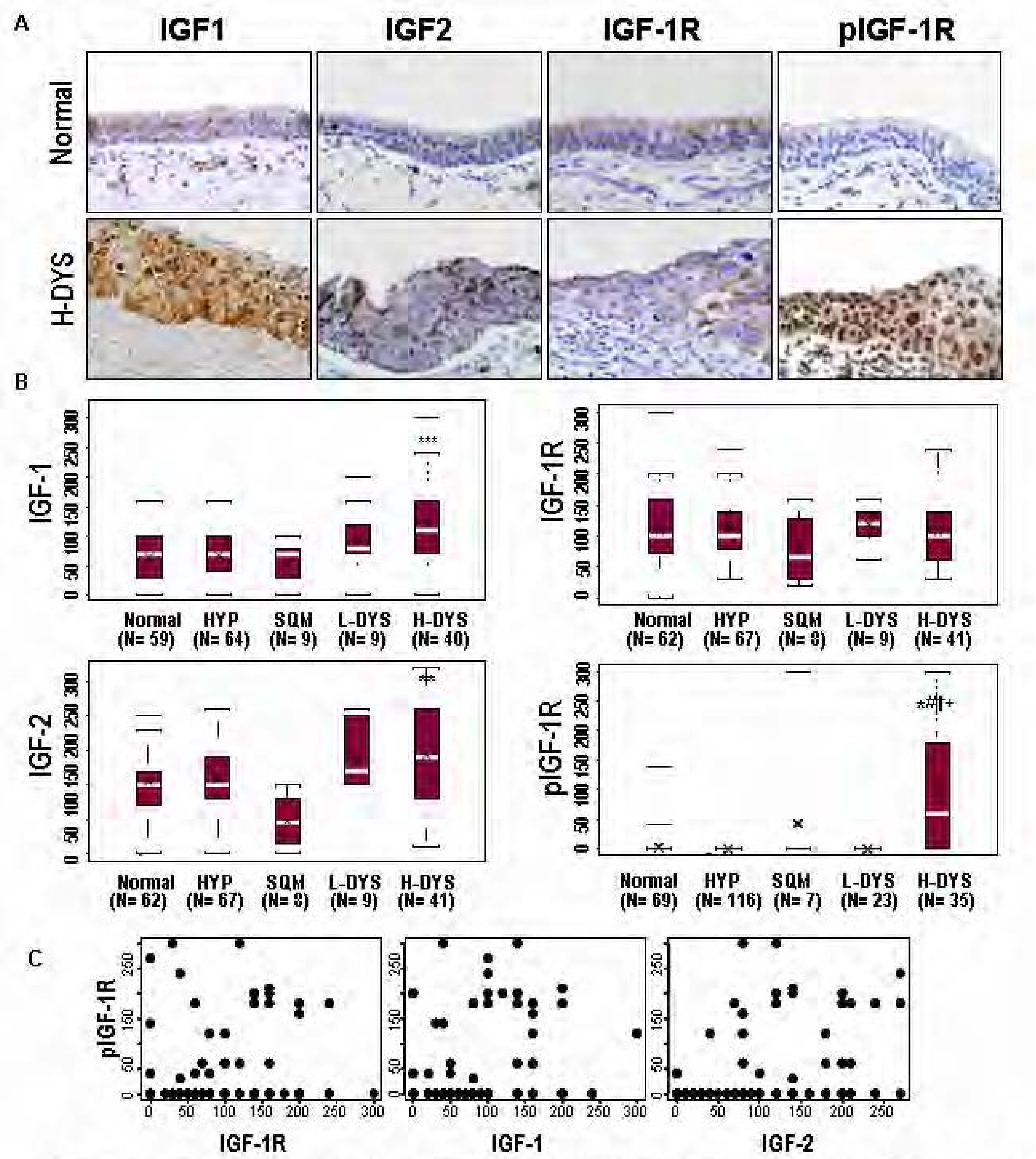

A
IGF1
IGF2
IGF-1R
pIGF-1R
Normal
H-DYS
B
IGF-1
0 50 100 150 200 250 300
***
Normal (N= 59)
HYP (N= 64)
SQM (N= 9)
L-DYS (N= 9)
H-DYS (N= 40)
IGF-1R
0 50 100 150 200 250 300
Normal (N= 62)
HYP (N= 67)
SQM (N= 8)
L-DYS (N= 9)
H-DYS (N= 41)
IGF-2
0 50 100 150 200 250 300
**
Normal (N= 62)
HYP (N= 67)
SQM (N= 8)
L-DYS (N= 9)
H-DYS (N= 41)
pIGF-1R
0 50 100 150 200 250 300
*#†+
Normal (N= 69)
HYP (N= 116)
SQM (N= 7)
L-DYS (N= 23)
H-DYS (N= 35)
C
pIGF-1R
0 50 150 250
0 50 100 150 200 250 300
IGF-1R
0 50 100 150 200 250 300
IGF-1
0 50 100 150 200 250
IGF-2

Figure 4. Immunohistochemical (IHC) scoring of insulin-like growth factor (IGF) -1 and IGF-2 in normal and preneoplastic lesions in patients with non–small cell lung cancer. A) Representative results of IHC staining of IGF signaling biomarkers in normal and high-grade dysplasia (H-DYS) specimens. B) Comparison of IHC scoring of IGF-1, IGF-2, IGF-1R, or pIGF-1R by histologic evaluation of normal, hyperplastic [HYP], squamous metaplastic [SQM] low-grade dysplasia [L-DYS] and high-grade dysplasia [H-DYS]. Box plots show the medians and the 25th and 75th percentiles; the bars give the ranges. Statistical significance of differences was determined with a repeated-measures model, and differences between groups were considered significant when *($P < 0.05$), ** ($P < 0.01$), *** ($P < 0.001$), statistically significant compared with normal tissue; # ($P < 0.05$), statistically significant compared with HYP; † ($P < 0.05$), statistically significant compared with SQM; + ($P < 0.05$), statistically significant compared with L-DYS. C) Linear regression analysis was performed to evaluate the significance of any association between the scoring of IGF-1R, IGF-1, or IGF-2 and the scoring of pIGF-1R in smokers and nonsmokers. The associations between p-IGFR and IGF-R, IGF-1 and IGF-2 were evaluated by using Spearman's rank correlation coefficient (0.08, 0.279, and 0.228, respectively) (The p values were 0.2373, 0.001, and 0.0015 respectively). The correlations were considered significant when $P < 0.05$.

Key Research Accomplishments

- IGF-1 and IGF-2 levels in bronchial tissue specimens containing high-grade dysplasia were significantly higher than in those containing normal epithelium, hyperplasia, and squamous metaplasia.
- Derivatives of human bronchial epithelial cell lines with activation mutation in *KRAS* (V12) or loss of p53 (genetic changes frequently observed during lung carcinogenesis) overexpressed IGF-1 and IGF-2. Tobacco carcinogen (TC) 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) enhanced transformed characteristics of these cells, which were significantly suppressed by inactivating IGFR.
- NNK and nicotine induced survivin protein synthesis in normal HBE cells by activating the Akt/mTOR pathway.

Conclusion

Our results demonstrate that airway epithelial cells produce IGFs in an autocrine manner, and these IGFs act jointly with TCs to stimulate lung carcinogenesis. Thus, the use of selective IGF-1R inhibitors may be a rational approach to controlling lung cancer. Blockade of the pathway effectively inhibits the tobacco-induced malignant transformation of HBE cells.

Core B: Biostatistics & Data Management Core

(Core Director: J. Jack Lee, Ph.D.)

Core Goals:

1. To provide statistical design, sample size/power calculations, and integrated, comprehensive analysis for each basic science, pre-clinical, and clinical study.
2. To develop a data management system that provides tracking, quality control, and integration of clinical, pathological, and basic science data. New database modules will be developed and integrated to the existing VITAL web-based database and with the clinical database from the Department of Thoracic/Head and Neck Medical Oncology.
3. To provide statistical and data management support for genomic and imaging studies including microarray, proteomics, protein antibody array, and spiral CT.
4. To develop and adapt innovative statistical methods pertinent to biomarker-integrated translational lung cancer studies.
5. To generate statistical reports for all projects.
6. To collaborate and assist all project investigators in the publication of scientific results.

Summary of Research Findings

The Biostatistics and Data Management Core has continued to work actively with all the VITAL Projects in their research efforts in the areas of biostatistical support and consulting in the clinical trial design, implementation, conduct, and analysis of experimental results.

Key Research Accomplishments

Our major effort in the fifth year was to continue to provide statistical and data management support for the “ReVITALization” of Project 1. A major database revamping effort has been accomplished to integrate the retrospectively collected data with the prospectively conducted clinical trial data. We have developed and continue to provide enhancement of a web-enabled database system to facilitate the research activities of the VANGUARD Trial. New database modules have been developed and integrated into the existing VITAL database and with the clinical database from the Department of Thoracic/Head and Neck Medical Oncology. The key activities of the database effort are summarized below.

- ReVITALization database has been developed to extend the VITAL database providing additional tissue clinical and pathological data repositories and tissue tracking.

- The SQL Server 2005 database and ASP.NET web application were used with VB.net language to provide the database infrastructure. Queries and SQL 2005 reports were implemented. Secure Socket Layer (SSL) and secured database password were applied to keep data transaction protected and confidential. Key components include:
  1) Clinical Module: The database's clinical module provides user friendly input interfaces for entering and viewing patients' clinical data and facilitates the patient search by medical record number (MRN), path number and name. The module contains the following web forms:
     - Patient Information
     - Social History (Alcohol and Smoking History)
     - Medical History
     - Other Malignancy
     - Treatments (Surgery, Chemotherapy, Radiotherapy and Other Treatments)
     - Clinical Staging
     - Follow-up

  2) The Revitalization's pathological module collects primary and metastasis general data, their histology, and tumor features, which can be used by our tool to automatically determine the cancer's staging. The tissue bank module of frozen (tissue, blood and pleural) and paraffin keep the tissue storage location and the concentration, volume and quality of the DNA, RNA and protein. The module contains the following web forms:
     - Primary and Metastasis Data (Diagnosis and Surgery Specimens)
     - Histology
     - Staging and Tumor Information: Cancer staging (TNM classification) is automatically determined by the system based on the tumor information provided.
     - Tissue Bank (Frozen Tissue and Paraffin)

  3) Reports: Several Excel reports are provided for the clinical and pathological modules.
     - Clinical Report
     - Pathological Report
     - Patient Report
     - Accession Report
     - General Information Report
     - Other Malignancy Report
     - Surgery Report
     - Chemotherapy Report
     - Radiotherapy Report
     - Other Treatment Report
     - Staging Report
     - Follow-up Report
     - Histology Diagnosis Report

4) Dictionaries: The database gives control for the users to conveniently add and update dictionaries; however, dictionary deletion is prohibited for users to prevent data loss. The database access is only allowed for the database administrators.

The Revitalization database application has been constantly updated to assist users in the data entering process. The retrospective data is checked and imported into the database system, while some new data has been entering by the users. The database is maintained and data is preserved by the database administrator. Selected screen shots are provided in the Appendix 1.

Major efforts are underway to support the ReVITALization effort. The database is fully functional and has captured 54 patients registered in the trial as of December 2008. The updating of clinical follow-up data for the retrospectively enrolled patients is in progress.

In addition, we have continued to work on statistical methods for evaluating interactions for combination therapy to determine whether the effects are synergistic, additive, or antagonistic. One manuscript has been published in *Biometrics* and one other is in press in *Statistics in Biopharmaceutical Research*. The related S-PLUS codes are available for download from http://biostatistics.mdanderson.org/SoftwareDownload/.

Conclusion

The Biostatistics and Data Management Core continues to provide the essential statistical and data management support for the entire VITAL project. As the project is coming to the end of the funding period, more efforts will be shift to data analysis, interpretation, and report writing.

Core C: Pathology and Specimen Procurement Core

(Core Director: Ignacio Wistuba, M.D.)

Aim 1. Develop and maintain a repository of tissue and other biologic specimens from patients enrolled on the clinical trials in Project 1.

Summary of Research Findings

In the past year, we have continued enrolling patients in the VITAL/Vanguard trials (Project 1). Up to December 2008, 53 patients have been enrolled and currently 31 patients have completed the baseline and 12-month bronchoscopy. From these patients, we have acquired, processed and banked a total of 1,451 specimens obtained during bronchoscopies (Table 1). In addition, resected specimens from the majority of lung cancer and head/neck tumors have been reviewed and banked.

Table 1. Summary of specimens collected and banked in the Pathology Core

Type of Specimen	Number
Sputum	**96**
Buccal Brush	**96**
Bronchial Brush	**581**
Bronchial Wash	**97**
Tissue Specimens	**581**
Total	**1,451**

Aim 2. Maintain a comprehensive database of tissue and specimen characteristics from patients enrolled in the clinical trials of Project 1, including pathologic characteristics of each specimen, inventory and distribution.

Summary of Research Findings

The Biostatistics Core has developed a web-based database that has been used to catalogue all the specimens obtained and banked in the Pathology Core and to report pathology diagnosis (see previous Core report). From the Vanguard patients, 870 cytological specimens and 581 bronchial biopsies have been tracked and inventoried using the Web-site database. As LIFE bronchoscopy biopsies have been performed on these patients, LIFE abnormalities have been banked in a database and will be correlated with histopathological features at a later time.

Aim 3. Provide comprehensive pathologic characterization of all tissues and other biologic specimens and assist in preparation and evaluation of studies involving these tissues.

Summary of Research Findings

We have processed and histopathologically diagnosed 581 tissue specimens from bronchoscopies (Table 2) with two H&E-stained tissue sections examined per bronchial biopsy. Although normal bronchial epithelium has been detected in at least one site examined in most subjects, a number of histopathological changes have been detected following tissue analysis. The most frequent abnormalities detected were goblet cell metaplasia (6.3%) and basal cell hyperplasia (17.9%). Of interest, squamous metaplasia (3.4%) and dysplasia (1.3%) were infrequently detected. No severe dysplasia, carcinoma *in situ,* or invasive carcinoma have been detected in these specimens. Histopathological analysis of the resected lung and head/neck cancer samples has been also been performed, and tissue blocks from all these samples are available for biomarker analysis.

Table 2. Summary of the 1,231 histopathology diagnoses made in 581 bronchial biopsies obtained from clinical trial (Project 1).

Diagnosis	N	%
No Tissue/Denuded Epithelium	**27**	**2.2**
Normal Epithelium	**848**	**68.9**
Goblet Cell Metaplasia (GCM)	**78**	**6.3**
Basal Cell Hyperplasia (BCH)	**208**	**16.9**
Combined GCM/BCH	**12**	**1.0**
Squamous Metaplasia	**42**	**3.4**
Angiogenic Squamous Dysplasia (ASD)	**2**	**0.2**
Mild Dysplasia	**8**	**0.6**
Moderate Dysplasia	**6**	**0.5**

Aim 4. Provide centralized immunohistochemistry and laser capture microdissection services, nucleic acid extractions and assistance with construction and evaluation of tissue arrays.

Summary of Research Findings

As reported previously, a centralized immunohistochemistry (IHC) laboratory is in place as part of the Pathology Core with manual and automated IHC techniques and *in situ* tissue-based methodologies, such as fluorescent *in situ* hybridization (FISH) and laser capture microdissection. Tissue microarray (TMA) construction is also in place, and a complete set of lung cancer specimens (N = 400, TMA set I) and preneoplastic lesions (preneoplasia TMA) has been constructed and utilized.

As reported previously, from a subset of early lung cancer surgically resected in our Institution from 1997 to 2002, we have identified and selected 504 histologically normal and abnormal bronchial, bronchiolar and alveolar epithelium specimens to be examined for molecular marker expression using IHC; all these specimens have been placed in TMAs. We have examined 31 IHC markers (Table 3), most of them related to potential targets for targeted-chemoprevention strategies. During the previous year, we have examined and identified 5 additional IHC markers in these preneoplasia specimens (Table 3).

Table 3. Summary of immunohistochemistry (IHC) markers examined in respiratory epithelium specimens placed in TMAs.

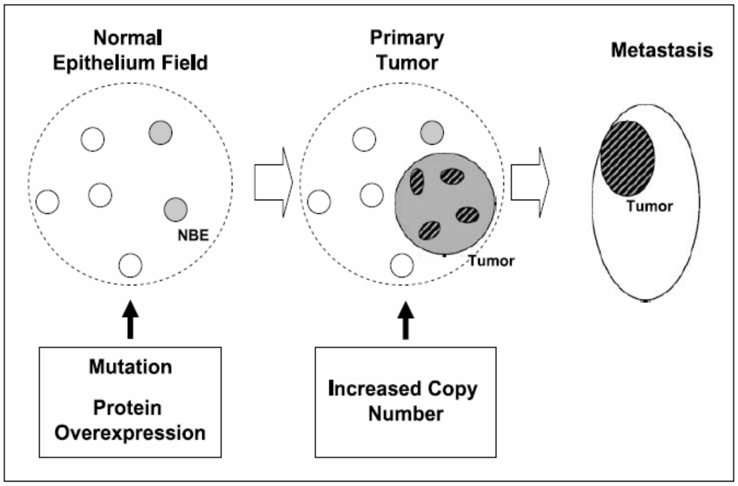
Immunohistochemistry Markers Examined in Respiratory Epithelium	
E-Cadherin (Project 2)	FGFR2 (Project 2)
Caspase-8 (Project 2)	Syndecan-1 (Project 2)
STAT-5 (Project 2)	Survivin (Project 4)
P70S6K (Project 2)	SCC-1 (Project 3)
NF-κB	SCC-2 (Project 3)
IRAK-1 (Project 2)	pCREB (Project 3)
IGF1 (Project 4)	CREB (Project 3)
IGFR1 (Project 4)	PTTG
p-IGFR1 (Project 4)*	Integrin- β6
IGF2 (Project 4)*	N-cadherin
EGFR (Project 4)	Vimentin
p-EGFR (Project 4)*	β-catenin
Caveolin-1	FUS-1
bFGF (Project 2)	MMP-9
FGFR1 (Project 2)	FEN1 (Project 2)*
UBE2C (Project 2)*	

* Markers examined during last year period

We examined the role of the expression of several pathways and multiple markers in the early pathogenesis of lung cancer using both tumor and preneoplasia TMAs. Last year, we published the characterization of the tumor suppressor gene *FUS-1* (1) and, in collaboration with Project 2, the transcription factor *CREB* (2) in NSCLC tumors and preneoplastic lesions. In addition, we submitted two articles for publication describing the role of estrogen and progesterone receptors (Raso et al., submitted to Clin Cancer Res) and epithelial-mesenchymal transition (EMT) phenomenon (Prudkin et al., submitted to Mod Pathol) in the early pathogenesis for lung cancer.

Last year, we published our data on the sequence of molecular events involving *EGFR* gene in the early pathogenesis and progression of *EGFR*-mutant lung adenocarcinomas (3) (Fig. 1). We demonstrated that *EGFR* mutations and protein overexpression are early phenomena in the pathogenesis of lung adenocarcinoma and that *EGFR* mutation precedes an increase in gene copy number. In *EGFR*-mutant adenocarcinoma metastases, the higher levels of EGFR overexpression and more homogeneously distributed high gene copy numbers suggest tumor progression. Our findings have important implications for the development of new strategies for targeted chemoprevention and therapy in lung adenocarcinoma using EGFR inhibitors. We have recently demonstrated that the *EGFR* T790M mutation, associated with resistance to treatment with EGFR small molecule inhibitors, is present in tumors not exposed to such treatment and may represent an early phenomenon in the pathogenesis of *EGFR* (T790M)-mutant tumors (4).

Figure 1. Proposed sequence of EGFR abnormalities occurring in the early pathogenesis and progression of EGFR-mutant lung adenocarcinomas. Normal bronchial epithelium (NBE) field, primary tumor, and metastasis sites. Small circles, NBE, which acquires EGFR mutations and EGFR protein (total and phosphorylated) overexpression (gray circles). In the primary tumor stage, the EGFR copy number increases (high polysomy and gene amplification) in small tumor foci (striped ovals). In the metastasis site, tumor cells show both EGFR mutation and high copy number throughout most of the lesion.



Finally, our studies on the role of the oncogene *TITF1* gene (5) in the pathogenesis of NSCLC showed that *TITF1* is amplified in both major types of NSCLC, adenocarcinomas and squamous cell carcinoma (18%) (Tang.et al., presented in the AACR Annual Meeting 2008, and manuscript in preparation), and in adenocarcinoma patients the combination of *EGFR* and *TITF1* amplification correlate with disease outcome (Tang et al., submitted to AACR Annual Meeting 2009, and manuscript in preparation).

Aim 5. Identify ~600 surgically resected tissue specimens from stages I/II NSCLC (tumor, normal and abnormal adjacent bronchial epithelium specimens) and their complete clinical and pathologic information.

Summary of Research Findings

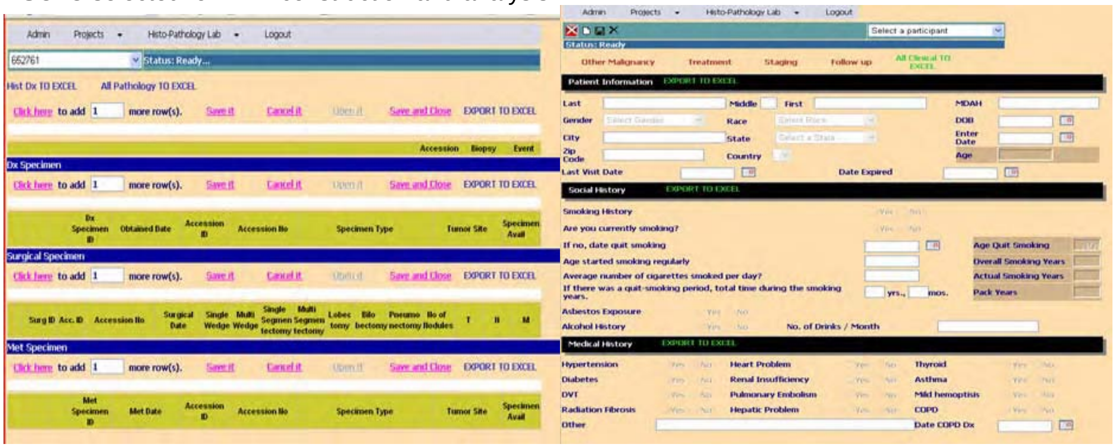
As part of the re-VITALization plan, we proposed to identify tumor and adjacent normal epithelium and preneoplastic lesions from 678 NSCLCs to be part of the TMA set II. We have completed the review of over 700 stages I to IIIA lung cancer surgically resected specimens in our Institution from January 2003 to December 2005 and obtained archival formalin-fixed paraffin-embedded specimens to identify tumor and adjacent normal or preneoplastic bronchial and bronchiolar epithelia. Of these, 678 cases have been fully characterized pathologically and TMA construction has been completed.

The detailed pathological information included the following: a) tumor histopathology: histology type (Table 4), percentage of histology sub-type for adenocarcinomas (bronchioloalveolar, acinar, papillary, solid and micropapillary), grade of differentiation, necrosis, fibrosis and inflammation; b) tumor gross features: type of surgical resection, lobule location, number of nodules, size, location (central vs. periphery), invasion (pleural, mediastinum, chest wall, vascular, neural and bronchial), and surgical margin. This detailed pathological analysis has been recently expanded to the 400 NSCLC cases included in the TMA set I.

Table 4. Histology diagnosis of the 678 NSCLC identified for TMA construction and marker analysis.

NSCLC Type	No. of tumors	%
Adenocarcinoma	412	60.8
Squamous cell carcinoma	212	31.3
Adenosquamous carcinoma	22	3.2
Large cell carcinoma	18	2.7
Sarcomatoid carcinoma	14	2.1
Total No. of Tumors	**678**	**100.0%**

Figure 2. Screens of the database utilized to enter detailed pathological and clinical data from the NSCLC selected for TMA construction and analysis.


Admin
Projects
Histo-Pathology Lab
Logout
652761
Status: Ready...
Hist Dx TO EXCEL
All Pathology TO EXCEL
Click here to add 1 more row(s).
Save it
Cancel it
Open it
Save and Close
EXPORT TO EXCEL
Accession
Biopsy
Event
Dx Specimen
Dx Specimen ID
Obtained Date
Accession ID
Accession No
Specimen Type
Tumor Site
Specimen Avail
Surgical Specimen
Surg ID
Acc. ID
Accession No
Surgical Date
Single Wedge
Multi Wedge
Single Segmentectomy
Multi Segmentectomy
Lobectomy
Bilobectomy
Pneumonectomy
No of Nodules
T
N
M
Met Specimen
Met Specimen ID
Met Date
Accession ID
Accession No
Specimen Type
Tumor Site
Specimen Avail
Select a participant
Status: Ready
Other Malignancy
Treatment
Staging
Follow up
All Clinical TO EXCEL
Patient Information
EXPORT TO EXCEL
Last
Middle
First
MDAH
Gender
Race
DOB
City
State
Enter Date
Zip Code
Country
Age
Last Visit Date
Date Expired
Social History
Smoking History
Are you currently smoking?
If no, date quit smoking
Age started smoking regularly
Average number of cigarettes smoked per day?
If there was a quit-smoking period, total time during the smoking years.
yrs.,
mos.
Age Quit Smoking
Overall Smoking Years
Actual Smoking Years
Pack Years
Asbestos Exposure
Alcohol History
No. of Drinks / Month
Medical History
Hypertension
Diabetes
DVT
Radiation Fibrosis
Other
Heart Problem
Renal Insufficiency
Pulmonary Embolism
Hepatic Problem
Thyroid
Asthma
Mild hemoptisis
COPD
Date COPD Dx

By reviewing the clinical charts, detailed demographic and clinical information has been obtained in all the cases, including follow-up information for adjuvant therapy, secondary tumor development, recurrence, metastasis pattern and survival. All the pathological and clinical data has been entered in database modules developed by the Biostatistic Core (Fig. 2). Currently, research personnel from Project 1 are initiating the process of contacting the alive patients to update the follow-up.

Aim 6. Examine over 40 biomarkers in those specimens by immunohistochemical (IHC) and tissue microarrays (TMAs).

Summary of Research Findings

TMA construction. Nineteen TMAs have been constructed using triplicate tissue cores (1 mm diameter) from 550 NSCLCs. Quality control of each tissue core (total 1.650 cores) has been completed and the TMAs histology sections have been sectioned for IHC analysis.

Immunohistochemistry (IHC) analysis. Using the experience obtained by analyzing nearly 130 markers in the NSCLC TMA set I, we have initiated the IHC analysis of the Re-VITALization TMA. Currently, in collaboration with Project 5 (P.I.: Dr. H-Y Lee), we are using IHC to examine the markers IGFR1 and p-IGFR1 in TMAs, and we are optimizing IGBP-3 and p-SRC. We have requested an unfunded, one-year extension to complete the IHC analysis of the remaining 36 markers. The IHC markers to be examined will be:

a) Markers obtained from among differentially expressed genes discovered by the analysis of normal human bronchial epithelial cells and premalignant, transformed and tumorigenic cell lines derived from them (Project 2, J. Minna, L. Mao, and R. Lotan) – 6 markers, some oncogene and TSGs abnormalities

b) Differentially expressed genes in abnormal areas of the bronchial epithelium identified LIFE-bronchoscopy (Project 3, P. Koo and W. Hittelman) – 6 markers

c) Death receptor (DR)-mediated apoptosis (Project 4, F. Khuri) – 6 markers

d) Akt and IGF/IGFR pathways markers (Project 5, H-Y Lee) – 8 markers

e) Markers obtained from the gene profiling analysis performed in 50-60 VANGUARD cases' specimens (Project 3, L. Mao) – 6 markers

Key Research Accomplishments

- Processed and diagnosed 1,450 bronchoscopy specimens from 53 patients enrolled in the Vanguard trial (Project 1).
- Completed the TMA construction and clinico-pathological data collection of 678 surgically resected NSCLC, and initiated the IHC analysis for outcome risk-model.
- Characterized the sequence of molecular events involved in the early pathogenesis and progression of *EGFR*-mutant lung adenocarcinomas.
- Characterized *TITF1* gene amplification in NSCLCs and demonstrated that gene amplification occurs in squamous cell carcinomas of the lung.

Conclusion

We have acquired and banked 1,450 specimens from bronchoscopies, resected specimens from lung cancer and head/neck tumor patients, and used the database developed by the Biostatistics Core to track and inventory bronchoscopy specimens and report histopathological features of 581 bronchial mucosa tissue specimens. We have examined by immunohistochemistry (IHC) 31 molecular markers in a large series (N=504) of respiratory epithelium specimens that have been used to examine several IHC markers related to VITAL-related projects. We have developed a working hypothesis for molecular events involved in the early pathogenesis of *EGFR*-mutant lung adenocarcinomas. We have completed the construction of tissue microarrays (TMA) and the collection of detailed clinico-pathological data from 678 NSCLCs, and initiated the molecular markers' analysis by IHC. During the one-year extension, we will complete the IHC analysis of this TMA set including 40 markers.

KEY RESEARCH ACCOMPLISHMENTS (IN SUMMARY)

Project 1: Biologic Approaches for Adjuvant Treatment of Aerodigestive Tract Cancer

- Enrolled 9 more patients in the Vanguard study with two additional patients pending, for a total of 55.
- Continued to collect patient clinical data and tissues for distribution to support research projects in the VITAL grant.
- 31 patients will have completed the 12-month mark/bronchoscopy.
- ReVITALization plan is being implemented with data entry into the pathology database related to the 500 archived tissue specimens, and updating of the clinical database for these patients.

Project 2: Identification of Biomarkers of Response to Chemoprevention Agents in Lung Epithelium

- Generated HBECS from over 45 different individuals and immortalized 10 peripheral small airway epithelial cells (HSAECs).
- Testing “preneoplastic” HSAECs with chemoprevention agents and performing microarray mRNA expression profiles and array-based CGH copy number profiles.
- Compared 104 cancer lines representing diverse tissue origins and discovered an unexpected frequency of activated genes in amplification 'hotspots.'
- Using qRT PCR technology on RNA isolated from microdissected tumors and normal lung epithelium from 30 patients, found nuclear hormone receptors whose expression differed between tumor and normal lung from the same patient, and between different lung cancers.
- Discovered that *BRF2* was amplified and overexpressed in preneoplastic carcinoma in situ (CIS) lesions in the airway epithelium of lung cancer patients
- Differentially expressed genes such as UBE2C, MCM2, FEN1, and BIRC5 are located within gene interaction networks significantly modulated between normal, immortalized, transformed and tumorigenic cells.
- The findings suggest that changes in MCM2, SFN, TPX2, FEN, and UBE2C genes are all important for the replication of immortalized, transformed and tumorigenic cells.
- UBE2C protein level is undetectable in normal appearing lung epithelium, but increases in aberrantly proliferating lesions and premalignant lesions. In addition, there was an increased expression in adenocarcinoma compared to normal epithelium.
- In lung cancer patients with stage I NSCLC, both overall survival and progression-free survival were statistically significantly lower in those cases where UBE2C expression was positive.
- Successfully amplified and hybridized 138 samples using Affymetrix ST1.0 Human gene chips.

Project 3: Premalignant Bronchial Epithelia: Molecular and Cellular Characterization of Lung Tumorigenesis

- Completed processing of 14 pair of patient samples and obtained the microarray data sets.
- Identified two distinct populations among patient samples, which has helped us to avoid the confounding effect of a mixed sample population and enabled us to find consensus in gene regulation patterns.
- Identified the cytokine-cytokine receptor pathway as one of the major signal pathways that may dictate the differences between the white-light or LIFE abnormal tissue and normal tissue, and could be critical in the early development of lung cancer.

- Located panels of genes that are unanimously regulated in each distinct sample cluster, and found that several of these genes coincide with genes previously reported to be important for tumorigenesis.
- Demonstrated that cells further along the multi-step tumorigenesis pathway tended to divide more frequently away from the basement membrane and their mitotic spindle orientation relative to the plane of the basal layer was highly variable.
- Confirmed that as cultures fashioned toward the tumor phenotype, they more frequently exhibited mitosis away from the basal layer and the orientation angle of the mitotic spindle (and degree of mitotic error) increased with distance from the basal layer.
- KG501, a small molecule inhibitor targeting CREB activity, and neutralizing antibodies against CXCL5 and CXCR2 blocked the migration of vascular endothelial cells induced by inflammatory cytokine.
- Demonstrated that EGFR ligands induce hyperplasia of bronchial cells cultured in an organotypic 3-dimensional culture method.
- Demonstrated that cyclin D1 upregulated by AP1 transcription factor plays a critical role in the hyperplasia of the bronchial epithelial cells and demonstrated that erlotinib blocks the hyperplasia.

Project 4: Modulation of Death Receptor-Mediated Apoptosis for Chemoprevention

- c-FLIP is overexpressed primarily in transformed malignant lung epithelial cell lines and may serve as a promising therapeutic target for prevention and therapy of lung cancer.
- The synthetic atypical retinoid SHetA2 induces apoptosis and enhances TRAIL-induced apoptosis involving up-regulation of DR5 and down-regulation of c-FLIP.
- CDDO-Me-induced up-regulation of death receptor 5 is coupled with endoplasmic reticulum (ER) stress via a CHOP-dependent mechanism involving JNK activation.
- Perifosine inhibits Akt/mTOR signaling and subsequently induces autophagy. Inhibition of this autophagy augments perifosine’s apoptosis-inducing activity, thus suggesting a therapeutic strategy of combining perifosine with an autophagy inhibitor.

Project 5: Molecular Strategies Targeting the AKT Signaling Pathway for Lung Cancer Chemoprevention and Therapy

- IGF-1 and IGF-2 levels in bronchial tissue specimens containing high-grade dysplasia were significantly higher than in those containing normal epithelium, hyperplasia, and squamous metaplasia.
- Derivatives of human bronchial epithelial cell lines with activation mutation in *KRAS* (V12) or loss of p53 (genetic changes frequently observed during lung carcinogenesis) overexpressed IGF-1 and IGF-2. Tobacco carcinogen (TC) 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) enhanced transformed characteristics of these cells, which were significantly suppressed by inactivating IGFR.
- NNK and nicotine induced survivin protein synthesis in normal HBE cells by activating the Akt/mTOR pathway.

Core B: Biostatistics & Data Management Core

- ReVITALization database has been developed to extend the VITAL database providing additional tissue clinical and pathological data repositories and tissue tracking.
- The SQL Server 2005 database and ASP.NET web application were used with VB.net language to provide the database infrastructure. Queries and SQL 2005 reports were implemented. Secure Socket Layer (SSL) and secured database password were applied to keep data transaction protected and confidential.

- Database is fully functional and has captured 54 patients registered in the trial as of December 2008.

Core C: Pathology and Specimen Procurement Core

- Processed and diagnosed 1,450 bronchoscopy specimens from 53 patients enrolled in the Vanguard trial (Project 1).
- Completed the TMA construction and clinico-pathological data collection of 678 surgically resected NSCLC, and initiated the IHC analysis for outcome risk-model.
- Characterized the sequence of molecular events involved in the early pathogenesis and progression of *EGFR*-mutant lung adenocarcinomas.
- Characterized *TITF1* gene amplification in NSCLCs and demonstrated that gene amplification occurs in squamous cell carcinomas of the lung.

REPORTABLE OUTCOMES

Publications *(attached in Appendix 2)***:**

Aggarwal S, Kim SW, Ryu SH, Chung WC, Koo JS. Growth Suppression of Lung Cancer Cells by Targeting Cyclic AMP Response Element-Binding Protein. *Cancer Research.* 2008; Feb 15;68(4):981-8.

Cucciarelli V, Hiser L, Smith H, Frankfurter A, Spano A, Correia JJ, Lobert S. b-Tubulin isotype classes II and V expression patterns in nonsmall cell lung carcinomas. *Cell Motility and the Cytoskeleton*. 2008; Aug;65(8):675-85.

Elrod HA, Sun SY. PPARgamma and Apoptosis in Cancer. *PPAR Research.* 2008;2008:704165.

Guha U, Chaerkady R, Marimuthu A, Patterson AS, Kashyap MK, Harsha HC, Sato M, Bader JS, Lash AE, Minna JD, Pandey A, Varmus HE. Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. *Proceedings of the National Academy of Sciences U S A.* 2008; Sep 16;105(37):14112-7.

Jin Q, Menter DG, Mao, L, Hong WK, Lee HY. Survivin expression in normal human brochial epithelial cells: an early and critical step in tumorigenesis induced by tobacco exposure. *Carcinogenesis.* 2008; Aug;29(8):1614-22.

Kong M, Lee JJ. A semiparametric response surface model for assessing drug interaction. *Biometrics*. 2008; Jun;64(2):396-405.

Kwei KA, Kim YH, Girard L, et al. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. *Oncogene.* 2008; Jun 5;27(25):3635-40.

Larsen JE, Spinola M, Gazdar AF, Minna JD An overview of the molecular biology of lung cancer. *In:* H. I. Pass, D. P. Carbone, D. H. Johnson, J. D. Minna, A. T. Turrisi III, and G. V. Scagliotti (eds.), *Lung Cancer: Principles and Practice*, 4th edition. Philadelphia: Lippincott Williams & Wilkins, 2009.

Lin YD, Chen S, Yue P, Zou W, Bendrook DM, Liu S, Le TC, Berlin KD, Khuri FR, Sun SY. CAAT/enhancer binding protein homologous protein-dependent end death receptor 5 induction is a major component of SHetA2-induced apoptosis in lung cancer cells. *Cancer Research*. 2008; Jul 1;68(13):5335-44.

Lin Y, Liu X, Yue P, Benbrook DM, Berlin KD, Khuri FR, Sun SY. Involvement of c-FLIP and survivin down-regulation in flexible heteroarotinoid-induced apoptosis and enhancement of TRAIL-initiated apoptosis in lung cancer cells. *Molecular Cancer Therapeutics.* 2008; Nov 7(11):3556-65.

Liu P, Vikis HG, Wang D, et al. Familial Aggregation of Common Sequence Variants on 15q24-25.1 in Lung Cancer. *Journal of the National Cancer Institute.* 2008; Sep 17;100(18):1326-30.

Lockwood WW, Chari R, Coe BP, et al. DNA amplification is a ubiquitous mechanism of oncogene activation in lung and other cancers. *Oncogene.* 2008;27:4615-24.

Prudkin L, Behrens C, Liu DD, Zhou X, Ozburn N, Bekele BN, et al. Loss and reduction of Fus1 protein expression is a frequent phenomenon in the pathogenesis of lung cancer. *Clinical Cancer Research.* 2008; Jan 1;14(1):41-7.

Prudkin L, Tang X, Wistuba II.Germ-line and somatic presentations of the EGFR T790M mutation in lung cancer. *Journal of Thoracic Oncology*. 2009 Jan; 4:139-4.

Raja SM, Chen S, Yue P, Acker TM, Lefkove B, Arbiser J, Khuri FR, Sun SY. The natural product honokiol preferentially inhibits c-FLIP and augments death receptor-induced apoptosis. *Molecular Cancer Therapeutics.* 2008; Jul;7(7):2212-23.

Seo HS, Liu DD, Bekele BN, Kim MK, Pisters K, Lippman SM, Wistuba II, Koo JS. Cyclic AMP response element-binding protein overexpression: a feature associated with negative prognosis in never smokers with non-small cell lung cancer. *Cancer Research.* 2008 Aug 1;68(15):6065-73.

Shivapurkar N, Stastny V, Xie Y, Prinsen C, Frenkel E, Czerniak B, Thunnissen FB, Minna JD, Gazdar AF. Differential methylation of a short CpG-rich sequence within exon 1 of TCF21 gene: a promising cancer biomarker assay. *Cancer Epidemiology Biomarkers Prevention.* 2008 Apr;17(4):995-1000.

Sun HX, Chung WC, Ryu SH, Ju ZL, Tran HT, Kim ES, Kurie JM, Koo JS. CREB- and NF-κB-Regulated CXC Chemokine Gene Expression in Lung Carcinogenesis. *Cancer Prevention Research.* 1(5):316-328, 2008.

Tang X, Varella-Garcia M, Xavier AC, Massarelli E, Ozburn N, Moran C, Wistuba I. Epidermal growth factor receptor abnormalities in the pathogenesis and progression of lung adenocarcinomas. *Cancer Prevention Research.* 2008; 1: 192-200.

Wistuba I, Meyerson M. Chromosomal Deletions and Progression of Premalignant Lesions. *Cancer Prevention Research.* 2008; 1;404-408.

Vikis H, Sato M, James M, et al. EGFR-T790M Is a Rare Lung Cancer Susceptibility Allele with Enhanced Kinase Activity. *Cancer Research.* 2007;67:4665-70.

Zhang L, Lee JJ, Tang H, Fan HT, Xiao L, Ren H, Kurie J, Morice RC, Hong WK, Mao L. Impact of smoking cessation on global gene expression in the bronchial epithelium of chronic smokers. *Cancer Prevention Research.* 2008; 1:112-118.

Zou W, Yue P, Khuri FR, Sun SY. Coupling of endoplasmic reticulum stress to CDDO-Me-induced up-regulation of death receptor 5 via a CHOP-dependent mechanism involving JNK activation. *Cancer Research.* 2008 Sep 15;68(18):7484-92.

Manuscripts submitted, in revision, or in review

Gazdar AF,Minna JD Deregulated EGFR signaling during lung cancer Progression: Mutations, amplicons and autocrine loops. *Cancer Prevention Research.* 2008;(In Press).

Ivanova MM, Mazhawidza W, Dougherty SM, Minna JD, Klinge CM Activity and intracellular location of estrogen receptors alpha and beta in human bronchial epithelial cells. *Molecular and Cellular Endocrinology.* 2008;(Submitted MCE-D-08-00399).

Kim YH, Girard L, Salari K, et al. Genomic and functional analysis identifies CRKL as an oncogene amplified in lung cancer. *Cancer Research.* 2008;(Submitted CAN-08-2785).

Lee JJ, Kong M. Confidence Interval of Interaction Index for Assessing Multiple Drug Interaction. *Statistics in Biopharmaceutical Research*. In Press, 2009.

Lockwood WW, Chari R, Coe BP, et al. BRF2 is a Lineage Specific Oncogene Amplified Early in Squamous Cell Lung Cancer Development. *Proceedings of the National Academy of Sciences U S A.* 2008;(Submitted PNAS 2008-11342).

McMahon K-A, Zajicek H, Li W-P, et al. SRBC/Cavin-3 Is A Caveolin Adapter Protein That Regulates Caveolae Function. *EMBO Journal.* 2009; (In Press).

Prudkin L., Liu D., Ozburn N., Sun M., Behrens C., Tang X., Brown K.C., Bekele B.N., Moran C., Wistuba I. Epithelial-Mesenchymal Transition in Non-Small Cell Lung Cancer Development and Progression. *Modern Pathology.* (Submitted).

Raso M.G., Behrens C., Liu S., Prudkin L., Ozburn N., Woods D.M., Tang X., Mehran R. J., Moran C., Lee J.J., Wistuba I. Immunohistochemical Expression of Estrogen and Progesterone Receptors Identifies a Subset of Non-small Cell Lung Carcinomas and Correlates with EGFR Mutation. *Clinical Cancer Research*. 2008 (In Press).

Woo-Young Kim, Quanri Jin, Seung-Hyun Oh, Edward S. Kim, Youn Joo Yang, Lei Feng, Carmen Behrens, Ludmila Prudkin, York E. Miller, J. Jack Lee, Scott M. Lippman, Waun Ki Hong Ignacio I. Wistuba, Ho-Young Lee. Elevated epithelial insulin-like growth factor expression is a risk factor for lung cancer development. *Cancer Research*. (In Revision).

You M, Wang D, Liu P, et al. Fine mapping of chromosome 6q23-25 region in familial lung cancer families reveals RGS17 as a likely candidate gene. *Clinical Cancer Research.* 2008 (In Press).

Abstracts (attached in Appendix 2)

Chen S, Fu L, Raja SM, Yue P, Peterson YK, Khuri FR, Sun S-Y. Differential roles of DR4, DR5 and c-FLIP in regulation of geranylgeranyltransferase I inhibitor-induced augmentation of tumor necrosis factor-related apoptosis-inducing ligand-induced apoptosis. Submitted to 2009 AACR.

Fan S, Li Y, Yue P, Khuri FR, Sun S-Y. The eIF4E/eIF4G interaction inhibitor 4EGI-1 augments TRAIL-induced apoptosis through DR5 induction and c-FLIP downregulation independent of inhibition of cap-dependent protein translation. Submitted to 2009 AACR.

Kim W-Y, Jin Q, Oh S-H, Kim E, Yang Y-J, Feng L, Behrens C, Prudkin L, Miller Y.E., Lee J.J., Lippman S, Hong W.K., Wistuba I, Lee H-Y. Elevated Epithelial Insulin-Like Growth Factor Expression is a Risk Factor for Lung Cancer Development. Submitted to 2009 AACR.

Lee JS, Ryu SH, Chung WC, Kang SM, and Koo JS. The Mechanism of EGF-Induced Hyperplasia in Normal Bronchial Epithelial Cells. Submitted to 2009 AACR.

Li X., Tang X., Behrens C, Dong W., Ozburn N., Woods D.M., Yin G., Hong W.K., Moran C., Wistuba I. STAT1 protein frequently overexpressed in non-small cell lung carcinoma. AACR Annual Meeting 2008. San Diego, California, April 2008.

Raso M.G., Behrens C., Liu S., Prudkin L., Woods D.M., Ozburn N. Moran C., Lee J.J., Wistuba I. Immunohistochemical expression of estrogen and progesterone receptors identifies a subset of non-small cell lung cancers and correlates with EGFR Mutations. AACR Annual Meeting 2008. San Diego, California, April 2008.

Tang X., Sun M., Behrens C., Prudkin L., Ozburn N., Gazdar A.F., Moran C., Varella-Garcia M., Wistuba I. TITF-1 gene amplification and protein expression pattern identify adenocarcinoma of lung with worse prognosis. AACR Annual Meeting 2008. San Diego, California, April 2008.

Tang X., Liu D., Behrens C., He D., Sun M., Rice D., Lee J.J., Hong W. K., Wistuba I. TITF-1 and EGFR gene copy variations are associated with prognosis for the patients with non-small cell lung cancer. Submitted to 2009 AACR.

CONCLUSIONS

Project 1: The ultimate goal is the development of a risk model for development of SPT and recurrence. This model will be developed utilizing the resources from both the Vanguard trial and our Lung Cancer Tissue Bank as a top priority for our program.

Project 2: Amplification is a far more common mechanism of oncogene activation than previously believed and specific regions of the genome are hotspots of amplification. These amplicons provide important new biomarkers that may be present in preneoplastic lesions and could be promising targets for the monitoring and chemoprevention of lung cancer. Nuclear hormone receptors (NRs) may provide prognostic value in lung cancer patients, and we would use these NRs expression profiles to hormonally manipulate lung cancer or as chemoprevention targets. Genetic alteration of BRF2 represents a novel mechanism of lung tumorigenesis through the increase of Pol III-mediated transcription in cancer.

Differentially expressed genes (e.g., UBE2C, MCM2, FEN1, and BIRC5) function in cell cycle control and display significant interactions with each other and with other known cancer-related genes in significant topological gene networks. Patient survival and progression-free survival were statistically significantly lower in those cases where UBE2C expression was positive.

Project 3: Cytokine-cytokine receptor pathway was identified as one of the major signal pathways that may dictate the differences between the white-light or LIFE abnormal tissue and normal tissue, and hence could play a critical role in the early development of lung cancer. Studies demonstrated that re-regulating the spatial location of mitotic events in the bronchial epithelium may decrease the rate of ongoing genetic instability and perhaps delay cancer onset in individuals with increased lung cancer risk.

KG501, a small molecule inhibitor targeting CREB activity, and neutralizing antibodies against CXCL5 and CXCR2 hold promise as preventive and/or a therapeutic strategy for lung cancer. EGFR ligands induce bronchial hyperplasia, which can be inhibited by erlotinib.

Project 4: Appropriate modulation of the extrinsic death receptor-mediated apoptotic pathway such as upregulation of DR5 and/or reduction of c-FLIP levels by small molecules may eliminate premalignant or malignant lung epithelial cells via promoting apoptotic cell death to achieve cancer chemopreventive and therapeutic goals.

The potential of the modulation of DR5 and c-FLIP as predictive biomarkers for certain drugs in the clinic warrants further investigation. Inhibition of autophagy may be a good strategy to augment certain anticancer agents' anticancer activity.

Project 5: Our results demonstrate that airway epithelial cells produce IGFs in an autocrine manner, and these IGFs act jointly with TCs to stimulate lung carcinogenesis. Thus, the use of selective IGF-1R inhibitors may be a rational approach to controlling lung cancer. Blockade of the pathway effectively inhibits the tobacco-induced malignant transformation of HBE cells.

Core B: The Biostatistics and Data Management Core continues to provide the essential statistical and data management support for the entire VITAL project. As the project is coming to the end of the funding period, more efforts will be shift to data analysis, interpretation, and report writing.

Core C: We have acquired and banked 1,450 specimens from bronchoscopies, resected specimens from lung cancer and head/neck tumor patients, and used the database developed by the Biostatistics Core to track and inventory bronchoscopy specimens and report histopathological features of 581 bronchial mucosa tissue specimens. We have examined by immunohistochemistry (IHC) 31 molecular markers in a large series (N=504) of respiratory epithelium specimens which have been used to examine several IHC markers related to VITAL-related projects. We have developed a working hypothesis for molecular events involved in the early pathogenesis of *EGFR*-mutant lung adenocarcinomas. We have completed the construction of tissue microarrays (TMA) and the collection of detailed clinico-pathological data from 678 NSCLCs, and initiated the molecular markers' analysis by IHC. During the one year extension, we will complete the IHC analysis of this TMA set including 40 markers.

References:

1. Adhami VM, Afaq F, Mukhtar H. Insulin-like growth factor-I axis as a pathway for cancer chemoprevention. *Clin Cancer Research.* 2006;12(19):5611-4.
2. Chun KH, Kosmeder JW, 2nd, Sun S, Pezzuto JM, Lotan R, Hong WK, et al. Effects of deguelin on the phosphatidylinositol 3-kinase/Akt pathway and apoptosis in premalignant human bronchial epithelial cells. *Journal of the National Cancer Institute.* 2003;95(4):291-302.
3. Elayadi AN, Samli KN, Prudkin L, et al. A Peptide Selected by Biopanning Identifies the Integrin {alpha}v{beta}6 as a Prognostic Biomarker for Nonsmall Cell Lung Cancer. *Cancer Research.* 2007;67:5889-95.
4. Garber, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proceedings of the National Academy of Sciences U S A.* 1998:13784-9, 2001.
5. Jin Q, Feng L, Behrens C, Bekele BN, Wistuba, II, Hong WK, et al. Implication of AMP-activated protein kinase and Akt-regulated survivin in lung cancer chemopreventive activities of deguelin. *Cancer Research.* 2007;67(24):11630-9.
6. Klein-Szanto AJ, Iizasa T, Momiki S, Garcia-Palazzo I, Caamano J, Metcalf R, et al. A tobacco-specific N-nitrosamine or cigarette smoke condensate causes neoplastic transformation of xenotransplanted human bronchial epithelial cells. *Proceedings of the National Academy of Sciences U S A.* 1992;89(15):6693-7.
7. Prudkin L, Behrens C, Liu DD, et al. Loss and reduction of FUS1 protein expression is a frequent phenomenon in the pathogenesis of lung cancer. *Clinical Cancer Research.* 2008;14:41-7.
8. Sell C, Dumenil G, Deveaud C, Miura M, Coppola D, DeAngelis T, et al. Effect of a null mutation of the insulin-like growth factor I receptor gene on growth and transformation of mouse embryo fibroblasts. *Molecular Cellular Biology.* 1994;14(6):3604-12.
8. Sato M, Vaughan MB, Girard L, Peyton M, Lee W, Shames DS, et al. Multiple oncogenic changes (K-RAS(V12), p53 knockdown, mutant EGFRs, p16 bypass, telomerase) are not sufficient to confer a full malignant phenotype on human bronchial epithelial cells. *Cancer Research*. 2006;66(4):2116-28.
9. Seo HS, Liu DD, Bekele BN, et al. Cyclic AMP response element-binding protein overexpression: a feature associated with negative prognosis in never smokers with non-small cell lung cancer. *Cancer Res*earch. 2008;68:6065-73.
10. Tang X, Varella-Garcia M, Xavier AC, et al. EGFR abnormalities in the pathogenesis and progression of lung adenocarcinomas. *Cancer Prevention Research.* 2008 Aug;1(3):192-200t.
11. Weir BA, Woo MS, Getz G, et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature*. 2007;450:893-8.

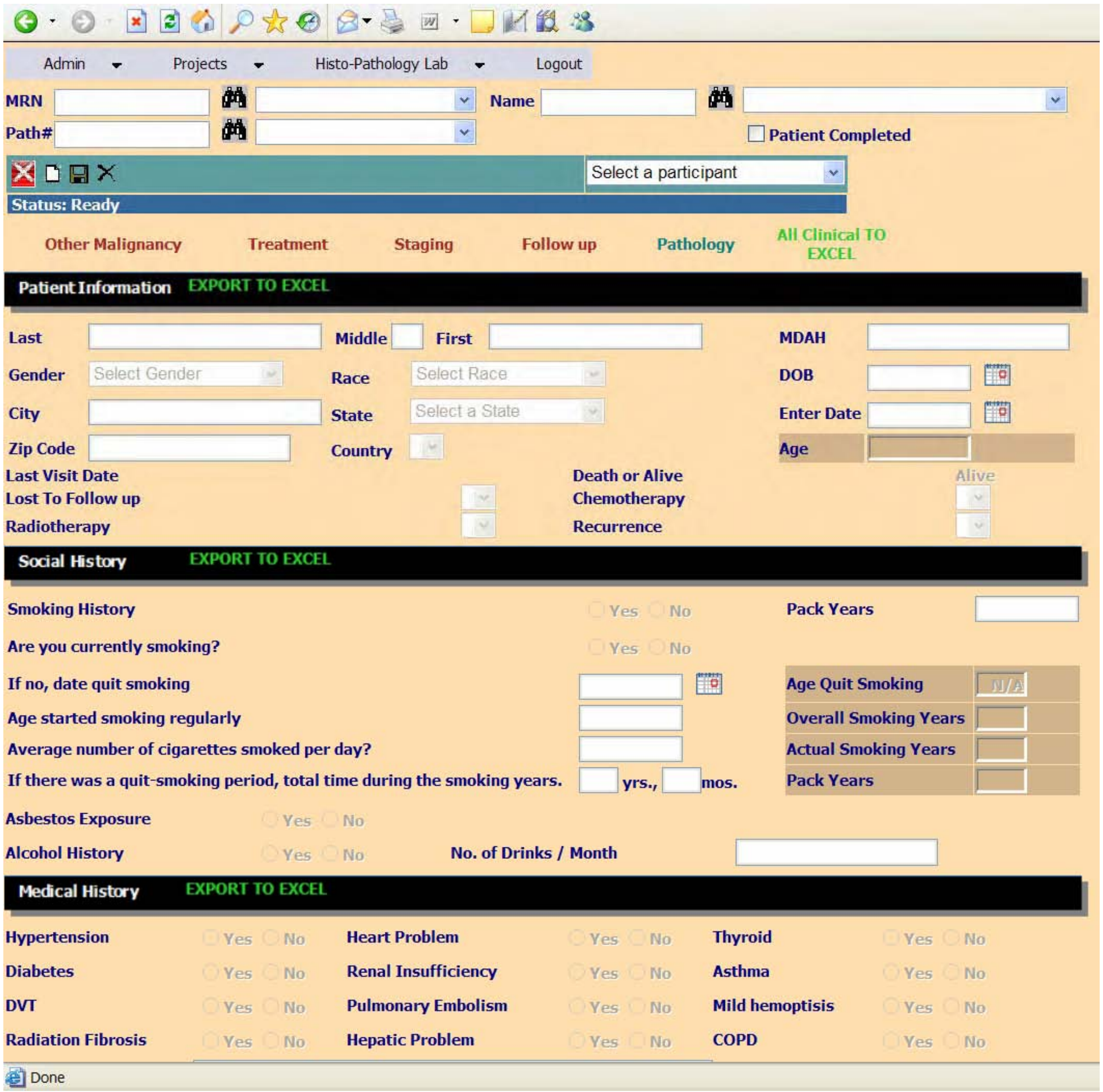
APPENDICES

APPENDIX 1

APPENDIX 2

Biostatistics Core Appendix

1) ReVITALization's clinical module: Patient Information, Social History, Medical History


Admin
Projects
Histo-Pathology Lab
Logout
Select a participant
Status: Ready
Other Malignancy
Treatment
Staging
Follow up
Pathology
All Clinical TO EXCEL
EXPORT TO EXCEL
Done

MRN
Path#
Name
☐ Patient Completed

Patient Information

Last
Middle
First
MDAH
Gender: Select Gender
Race: Select Race
DOB
City
State: Select a State
Enter Date
Zip Code
Country
Age
Last Visit Date
Lost To Follow up
Radiotherapy
Death or Alive: Alive
Chemotherapy
Recurrence

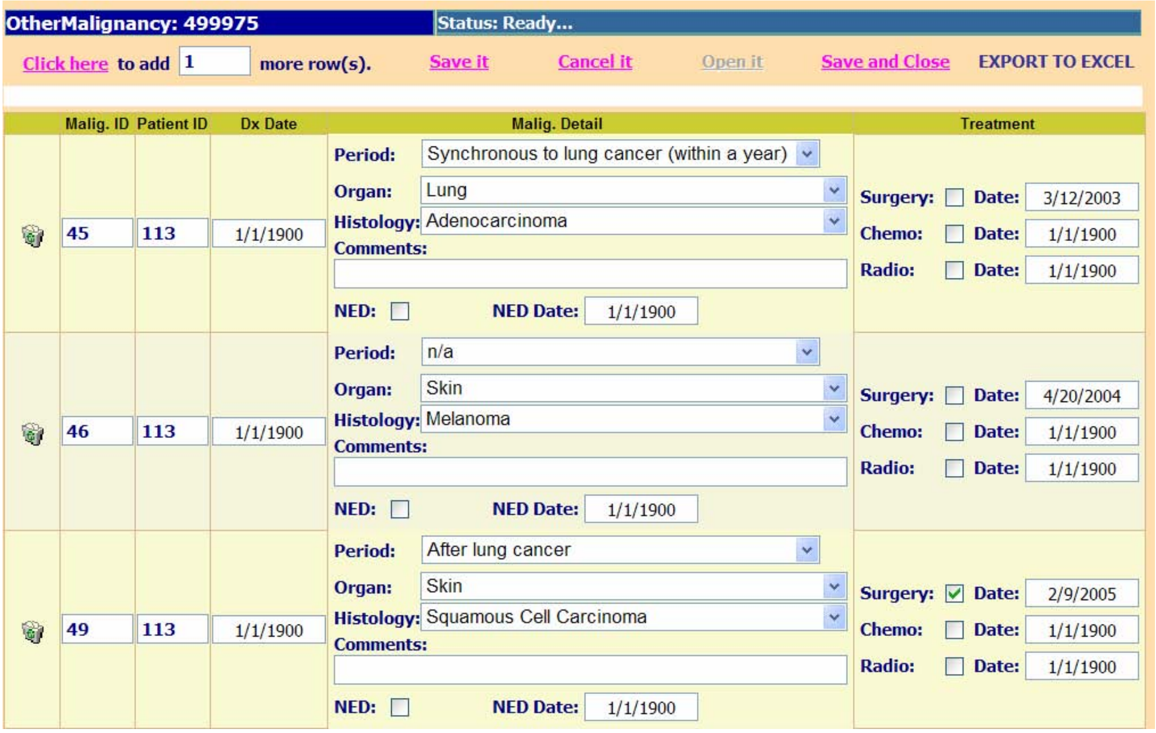
Social History

Smoking History ○ Yes ○ No
Pack Years
Are you currently smoking? ○ Yes ○ No
If no, date quit smoking
Age Quit Smoking: N/A
Age started smoking regularly
Overall Smoking Years
Average number of cigarettes smoked per day?
Actual Smoking Years
If there was a quit-smoking period, total time during the smoking years. yrs., mos.
Pack Years
Asbestos Exposure ○ Yes ○ No
Alcohol History ○ Yes ○ No
No. of Drinks / Month

Medical History

Hypertension	○ Yes ○ No	Heart Problem	○ Yes ○ No	Thyroid	○ Yes ○ No
Diabetes	○ Yes ○ No	Renal Insufficiency	○ Yes ○ No	Asthma	○ Yes ○ No
DVT	○ Yes ○ No	Pulmonary Embolism	○ Yes ○ No	Mild hemoptisis	○ Yes ○ No
Radiation Fibrosis	○ Yes ○ No	Hepatic Problem	○ Yes ○ No	COPD	○ Yes ○ No

2) ReVITALization’s clinical module: Other Malignancy

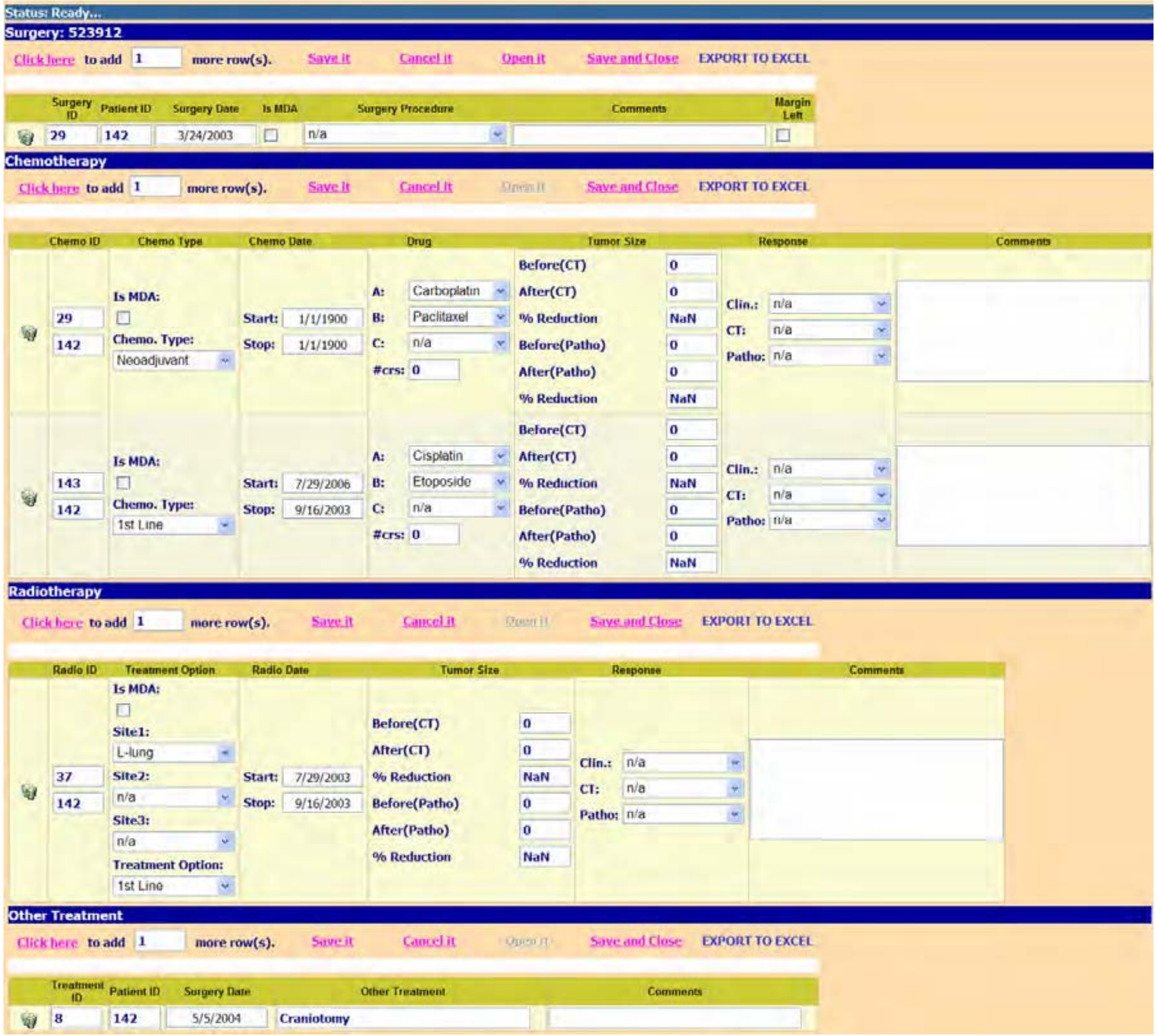


OtherMalignancy: 499975 Status: Ready...

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL

	Malig. ID	Patient ID	Dx Date	Malig. Detail	Treatment
	45	113	1/1/1900	**Period:** Synchronous to lung cancer (within a year) **Organ:** Lung **Histology:** Adenocarcinoma **Comments:** **NED:** ☐ **NED Date:** 1/1/1900	**Surgery:** ☐ **Date:** 3/12/2003 **Chemo:** ☐ **Date:** 1/1/1900 **Radio:** ☐ **Date:** 1/1/1900
	46	113	1/1/1900	**Period:** n/a **Organ:** Skin **Histology:** Melanoma **Comments:** **NED:** ☐ **NED Date:** 1/1/1900	**Surgery:** ☐ **Date:** 4/20/2004 **Chemo:** ☐ **Date:** 1/1/1900 **Radio:** ☐ **Date:** 1/1/1900
	49	113	1/1/1900	**Period:** After lung cancer **Organ:** Skin **Histology:** Squamous Cell Carcinoma **Comments:** **NED:** ☐ **NED Date:** 1/1/1900	**Surgery:** ☑ **Date:** 2/9/2005 **Chemo:** ☐ **Date:** 1/1/1900 **Radio:** ☐ **Date:** 1/1/1900

3) ReVITALization's clinical module: Treatment: Surgery, Chemotherapy, Radiotherapy and Other Treatments.



Surgery: 523912

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL

	Surgery ID	Patient ID	Surgery Date	Is MDA	Surgery Procedure	Comments	Margin Left
	29	142	3/24/2003	☐	n/a		☐

Chemotherapy

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL

	Chemo ID	Chemo Type	Chemo Date	Drug	Tumor Size	Response	Comments
	29 142	Is MDA: ☐ Chemo. Type: Neoadjuvant	Start: 1/1/1900 Stop: 1/1/1900	A: Carboplatin B: Paclitaxel C: n/a #crs: 0	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	
	143 142	Is MDA: ☐ Chemo. Type: 1st Line	Start: 7/29/2006 Stop: 9/16/2003	A: Cisplatin B: Etoposide C: n/a #crs: 0	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	

Radiotherapy

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL

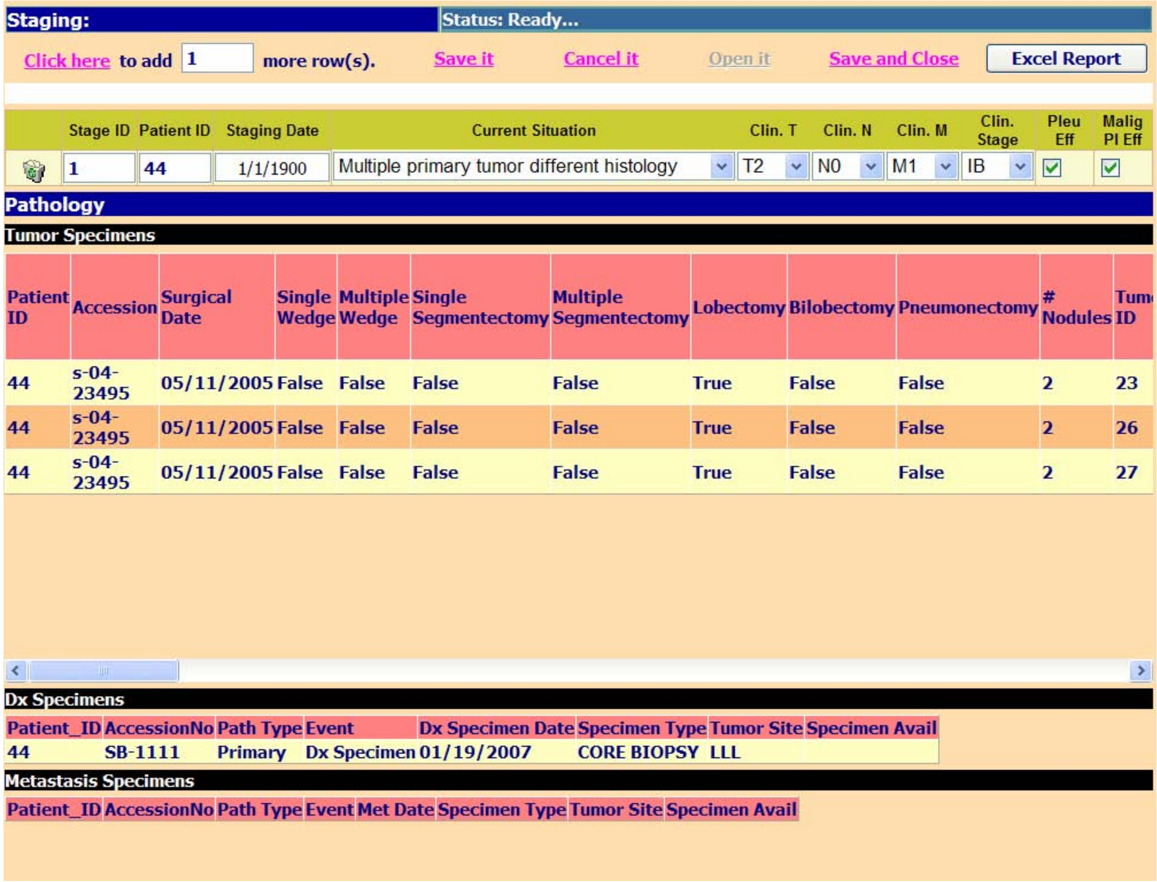
	Radio ID	Treatment Option	Radio Date	Tumor Size	Response	Comments
	37 142	Is MDA: ☐ Site1: L-lung Site2: n/a Site3: n/a Treatment Option: 1st Line	Start: 7/29/2003 Stop: 9/16/2003	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	

Other Treatment

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL

	Treatment ID	Patient ID	Surgery Date	Other Treatment	Comments
	8	142	5/5/2004	Craniotomy	

4) ReVITALization's clinical module: Staging



	Stage ID	Patient ID	Staging Date	Current Situation	Clin. T	Clin. N	Clin. M	Clin. Stage	Pleu Eff	Malig Pl Eff
	1	44	1/1/1900	Multiple primary tumor different histology	T2	N0	M1	IB	☑	☑

Pathology

Tumor Specimens

Patient ID	Accession	Surgical Date	Single Wedge	Multiple Wedge	Single Segmentectomy	Multiple Segmentectomy	Lobectomy	Bilobectomy	Pneumonectomy	# Nodules	Tum ID
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	23
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	26
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	27

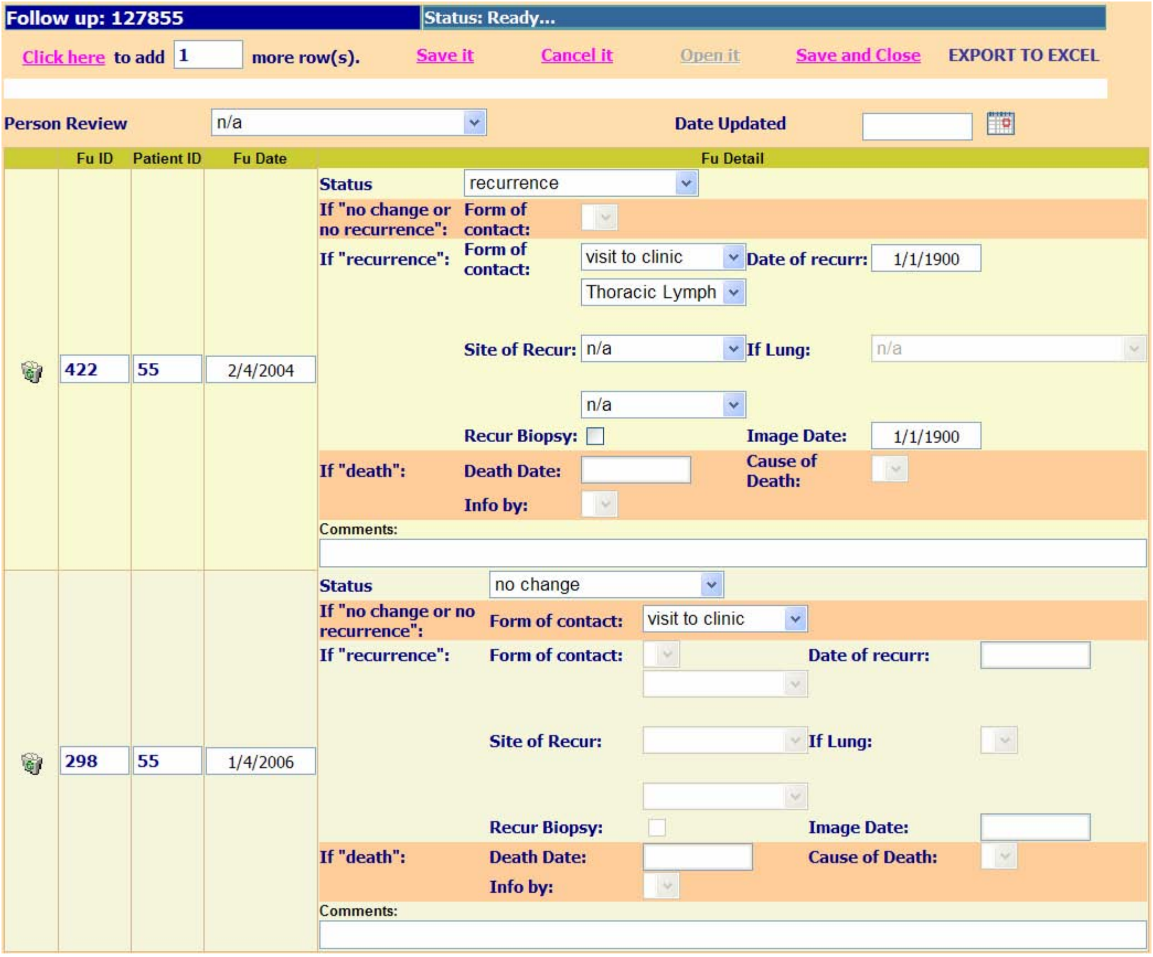
Dx Specimens

Patient_ID	AccessionNo	Path Type	Event	Dx Specimen Date	Specimen Type	Tumor Site	Specimen Avail
44	SB-1111	Primary	Dx Specimen	01/19/2007	CORE BIOPSY	LLL	

Metastasis Specimens

Patient_ID	AccessionNo	Path Type	Event	Met Date	Specimen Type	Tumor Site	Specimen Avail

5) ReVITALization's clinical module: Follow up



Follow up: 127855

Status: Ready...

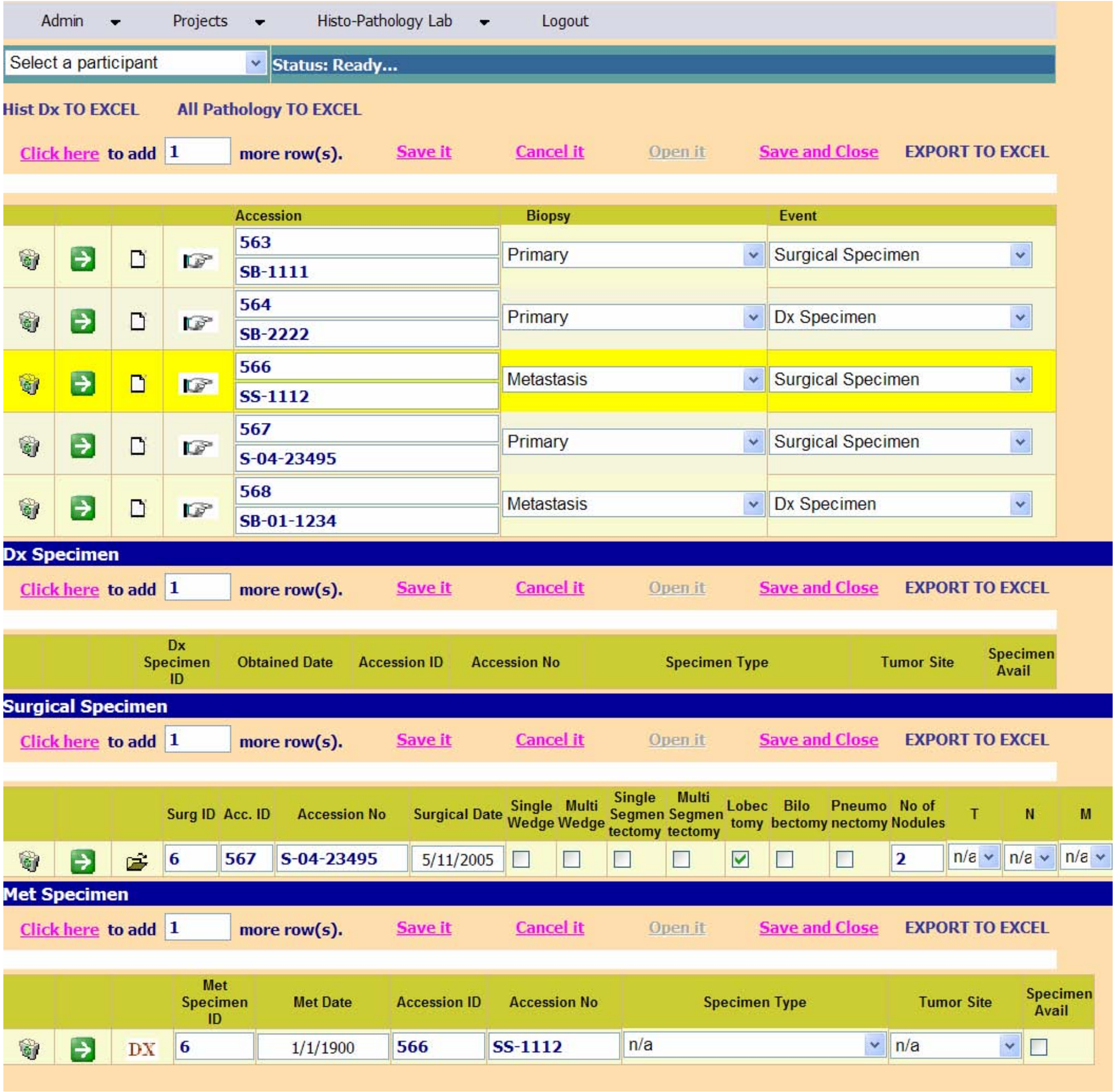
Person Review n/a

Date Updated

Fu ID	Patient ID	Fu Date	Fu Detail
422	55	2/4/2004	**Status** recurrence **If "no change or no recurrence":** **Form of contact:** **If "recurrence":** **Form of contact:** visit to clinic **Date of recurr:** 1/1/1900 Thoracic Lymph **Site of Recur:** n/a **If Lung:** n/a n/a **Recur Biopsy:** ☐ **Image Date:** 1/1/1900 **If "death":** **Death Date:** **Cause of Death:** **Info by:** Comments:
298	55	1/4/2006	**Status** no change **If "no change or no recurrence":** **Form of contact:** visit to clinic **If "recurrence":** **Form of contact:** **Date of recurr:** **Site of Recur:** **If Lung:** **Recur Biopsy:** ☐ **Image Date:** **If "death":** **Death Date:** **Cause of Death:** **Info by:** Comments:

6) ReVITALization's pathological module: Tissue Pathological Data

- Primary Dx specimen
- Primary Surgical Specimen
- Metastasis Dx Specimen
- Metastasis Surgical specimen


Admin
Projects
Histo-Pathology Lab
Logout
Select a participant
Hist Dx TO EXCEL
All Pathology TO EXCEL
Click here to add 1 more row(s).
Save it
Cancel it
Open it
Save and Close
EXPORT TO EXCEL

Status: Ready...

Accession	Biopsy	Event
563 SB-1111	Primary	Surgical Specimen
564 SB-2222	Primary	Dx Specimen
566 SS-1112	Metastasis	Surgical Specimen
567 S-04-23495	Primary	Surgical Specimen
568 SB-01-1234	Metastasis	Dx Specimen

Dx Specimen

Dx Specimen ID	Obtained Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail

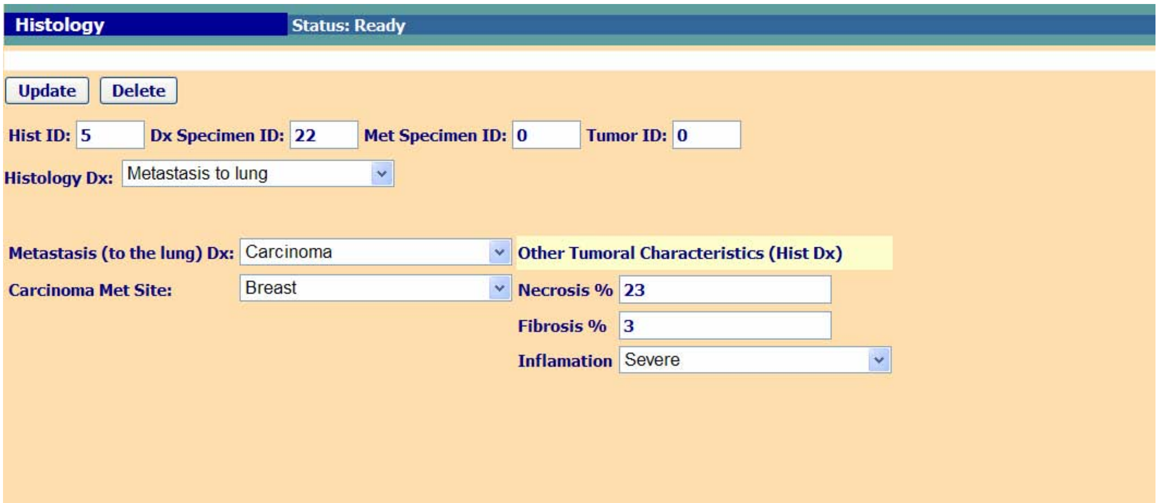
Surgical Specimen

Surg ID	Acc. ID	Accession No	Surgical Date	Single Wedge	Multi Wedge	Single Segmentectomy	Multi Segmentectomy	Lobectomy	Bilobectomy	Pneumonectomy	No of Nodules	T	N	M
6	567	S-04-23495	5/11/2005	☐	☐	☐	☐	☑	☐	☐	2	n/a	n/a	n/a

Met Specimen

	Met Specimen ID	Met Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail
DX	6	1/1/1900	566	SS-1112	n/a	n/a	☐

7) ReVITALization's pathological module: Histology



Hist ID: 5 Dx Specimen ID: 22 Met Specimen ID: 0 Tumor ID: 0

Histology Dx: Metastasis to lung

Metastasis (to the lung) Dx: Carcinoma

Carcinoma Met Site: Breast

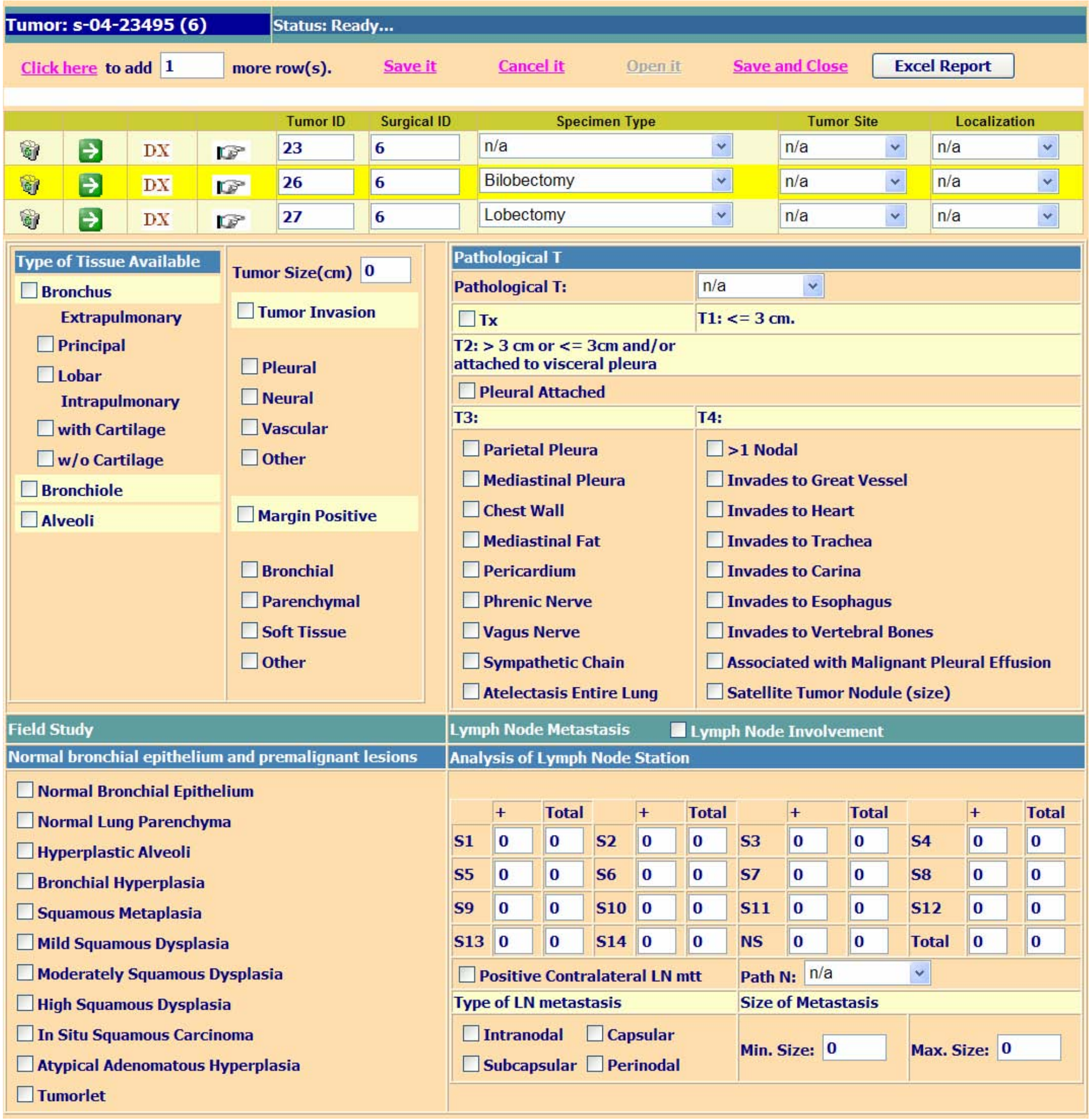
Other Tumoral Characteristics (Hist Dx)

Necrosis % 23

Fibrosis % 3

Inflamation Severe

8) ReVITALization's pathological module: Staging and Tumor Information



Tumor: s-04-23495 (6) Status: Ready...

Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report

				Tumor ID	Surgical ID	Specimen Type	Tumor Site	Localization
		DX		23	6	n/a	n/a	n/a
		DX		26	6	Bilobectomy	n/a	n/a
		DX		27	6	Lobectomy	n/a	n/a

Type of Tissue Available

- ☐ Bronchus
  - Extrapulmonary
    - ☐ Principal
    - ☐ Lobar
  - Intrapulmonary
    - ☐ with Cartilage
    - ☐ w/o Cartilage
- ☐ Bronchiole
- ☐ Alveoli

Tumor Size(cm) 0

- ☐ Tumor Invasion
  - ☐ Pleural
  - ☐ Neural
  - ☐ Vascular
  - ☐ Other
- ☐ Margin Positive
  - ☐ Bronchial
  - ☐ Parenchymal
  - ☐ Soft Tissue
  - ☐ Other

Pathological T

Pathological T: n/a

☐ Tx

T1: <= 3 cm.

T2: > 3 cm or <= 3cm and/or attached to visceral pleura

☐ Pleural Attached

T3:	T4:
☐ Parietal Pleura	☐ >1 Nodal
☐ Mediastinal Pleura	☐ Invades to Great Vessel
☐ Chest Wall	☐ Invades to Heart
☐ Mediastinal Fat	☐ Invades to Trachea
☐ Pericardium	☐ Invades to Carina
☐ Phrenic Nerve	☐ Invades to Esophagus
☐ Vagus Nerve	☐ Invades to Vertebral Bones
☐ Sympathetic Chain	☐ Associated with Malignant Pleural Effusion
☐ Atelectasis Entire Lung	☐ Satellite Tumor Nodule (size)

Field Study

Normal bronchial epithelium and premalignant lesions

- ☐ Normal Bronchial Epithelium
- ☐ Normal Lung Parenchyma
- ☐ Hyperplastic Alveoli
- ☐ Bronchial Hyperplasia
- ☐ Squamous Metaplasia
- ☐ Mild Squamous Dysplasia
- ☐ Moderately Squamous Dysplasia
- ☐ High Squamous Dysplasia
- ☐ In Situ Squamous Carcinoma
- ☐ Atypical Adenomatous Hyperplasia
- ☐ Tumorlet

Lymph Node Metastasis ☐ Lymph Node Involvement

Analysis of Lymph Node Station

	+	Total		+	Total		+	Total		+	Total
S1	0	0	S2	0	0	S3	0	0	S4	0	0
S5	0	0	S6	0	0	S7	0	0	S8	0	0
S9	0	0	S10	0	0	S11	0	0	S12	0	0
S13	0	0	S14	0	0	NS	0	0	Total	0	0

☐ Positive Contralateral LN mtt

Path N: n/a

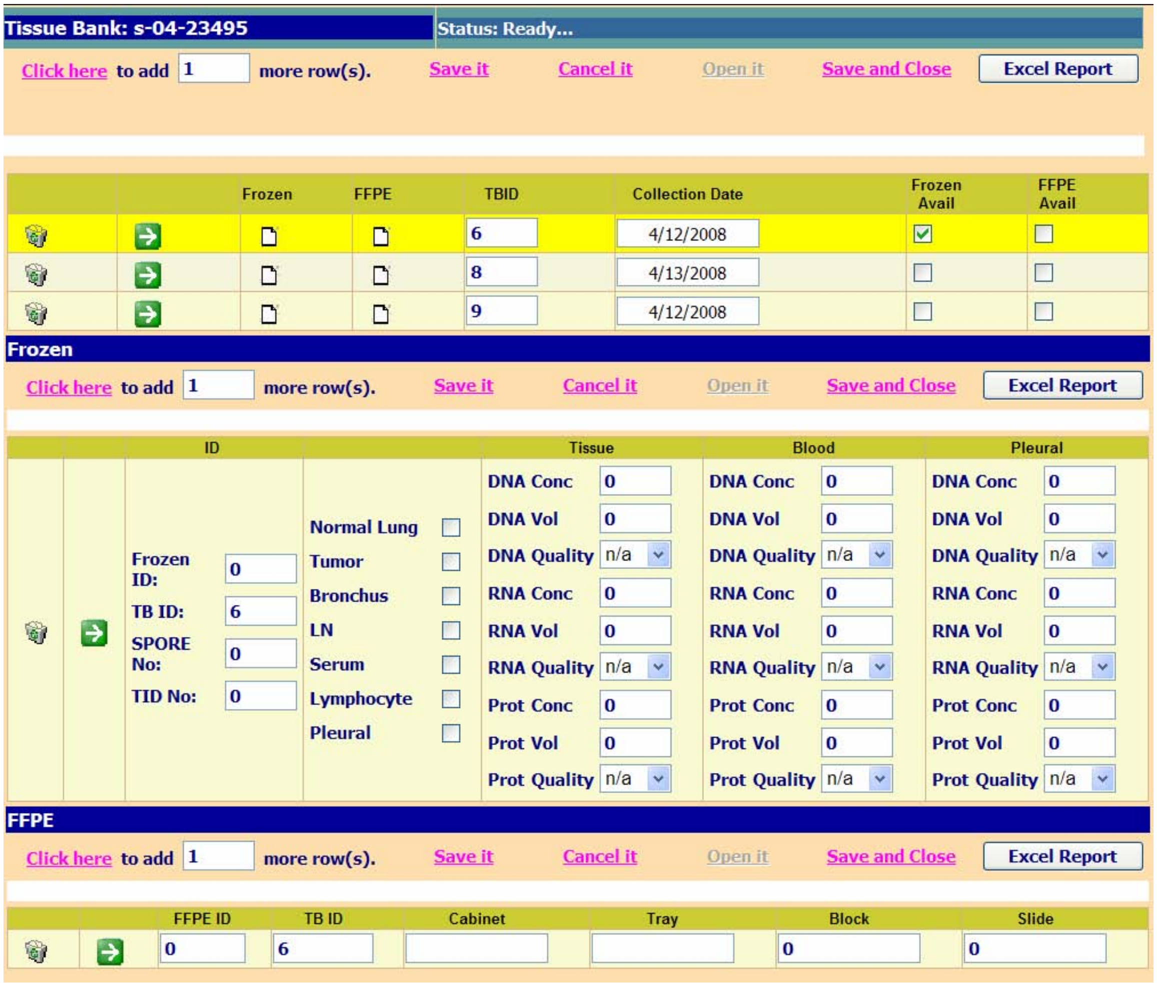
Type of LN metastasis

☐ Intranodal ☐ Capsular
☐ Subcapsular ☐ Perinodal

Size of Metastasis

Min. Size: 0
Max. Size: 0

9) ReVITALization's pathological module: Tissue Bank (Frozen and Paraffin)


Click here to add 1 more row(s).
Save it
Cancel it
Open it
Save and Close
Excel Report

Tissue Bank: s-04-23495

Status: Ready...

		Frozen	FFPE	TBID	Collection Date	Frozen Avail	FFPE Avail
				6	4/12/2008	☑	☐
				8	4/13/2008	☐	☐
				9	4/12/2008	☐	☐

Frozen

Click here to add 1 more row(s).
Save it
Cancel it
Open it
Save and Close
Excel Report

ID

Frozen ID: 0
TB ID: 6
SPORE No: 0
TID No: 0

Normal Lung ☐
Tumor ☐
Bronchus ☐
LN ☐
Serum ☐
Lymphocyte ☐
Pleural ☐

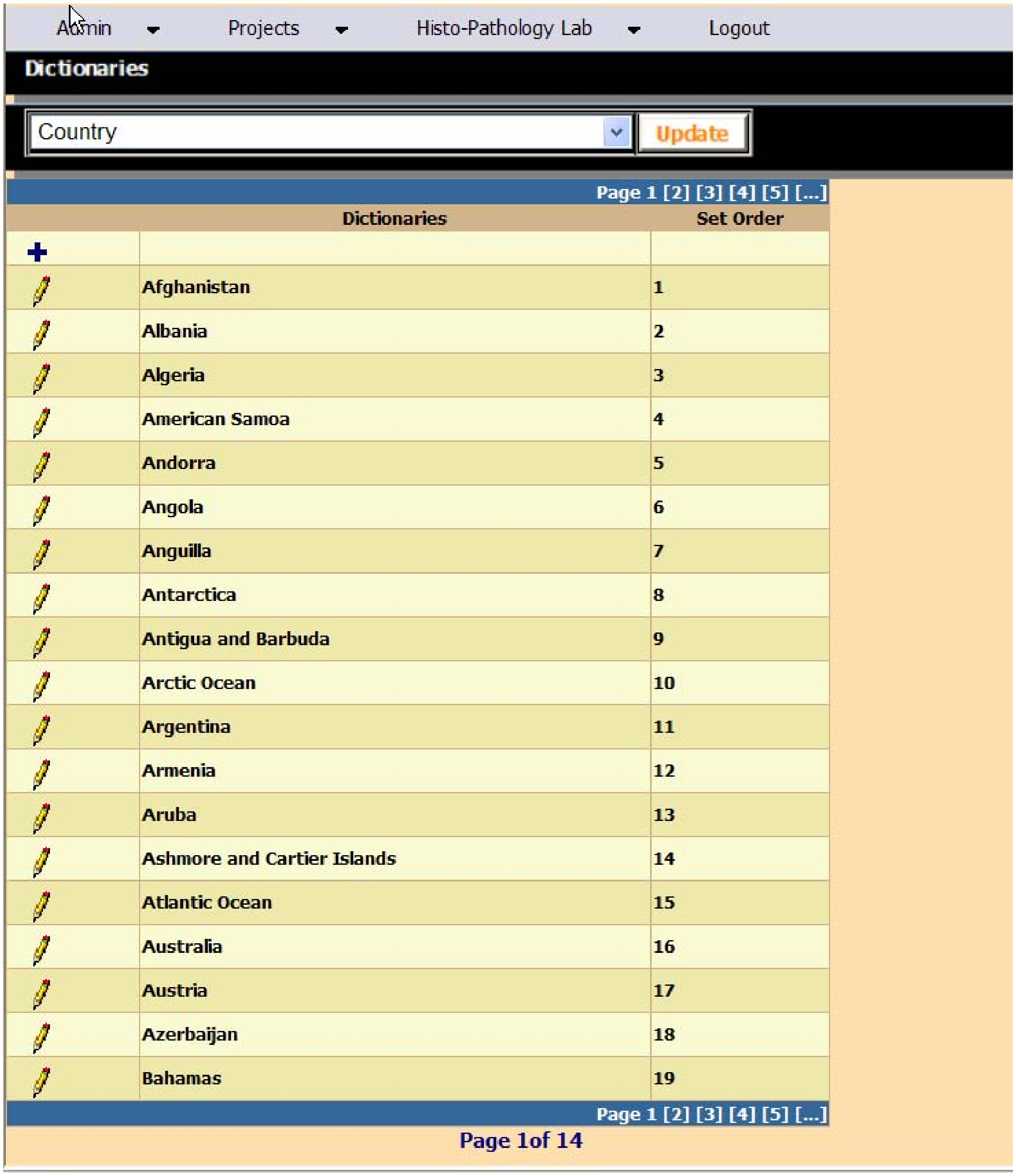
	Tissue	Blood	Pleural
DNA Conc	0	0	0
DNA Vol	0	0	0
DNA Quality	n/a	n/a	n/a
RNA Conc	0	0	0
RNA Vol	0	0	0
RNA Quality	n/a	n/a	n/a
Prot Conc	0	0	0
Prot Vol	0	0	0
Prot Quality	n/a	n/a	n/a

FFPE

Click here to add 1 more row(s).
Save it
Cancel it
Open it
Save and Close
Excel Report

		FFPE ID	TB ID	Cabinet	Tray	Block	Slide
		0	6			0	0

10) ReVITALization's dictionary module


Admin
Projects
Histo-Pathology Lab
Logout
Dictionaries
Country
Update

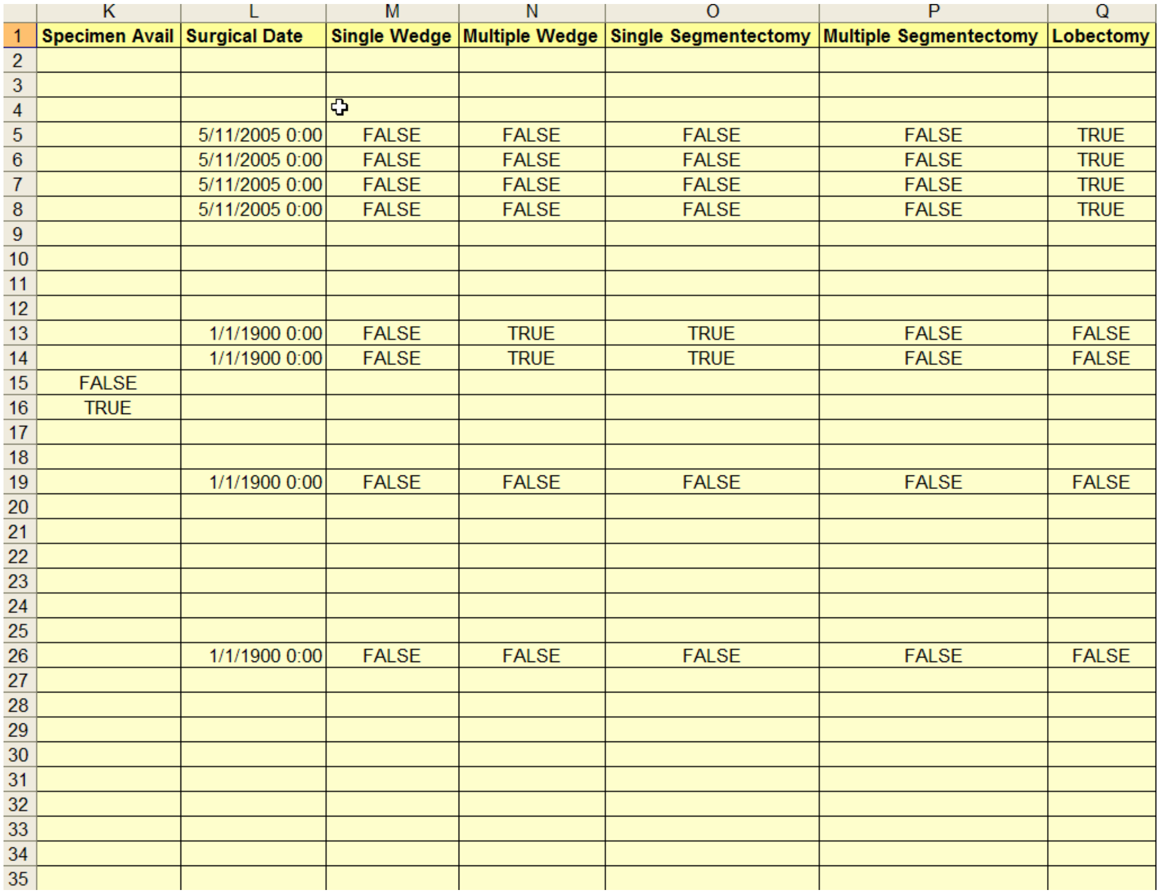
Page 1 [2] [3] [4] [5] [...]

	Dictionaries	Set Order
+		
✏	Afghanistan	1
✏	Albania	2
✏	Algeria	3
✏	American Samoa	4
✏	Andorra	5
✏	Angola	6
✏	Anguilla	7
✏	Antarctica	8
✏	Antigua and Barbuda	9
✏	Arctic Ocean	10
✏	Argentina	11
✏	Armenia	12
✏	Aruba	13
✏	Ashmore and Cartier Islands	14
✏	Atlantic Ocean	15
✏	Australia	16
✏	Austria	17
✏	Azerbaijan	18
✏	Bahamas	19

Page 1 [2] [3] [4] [5] [...]

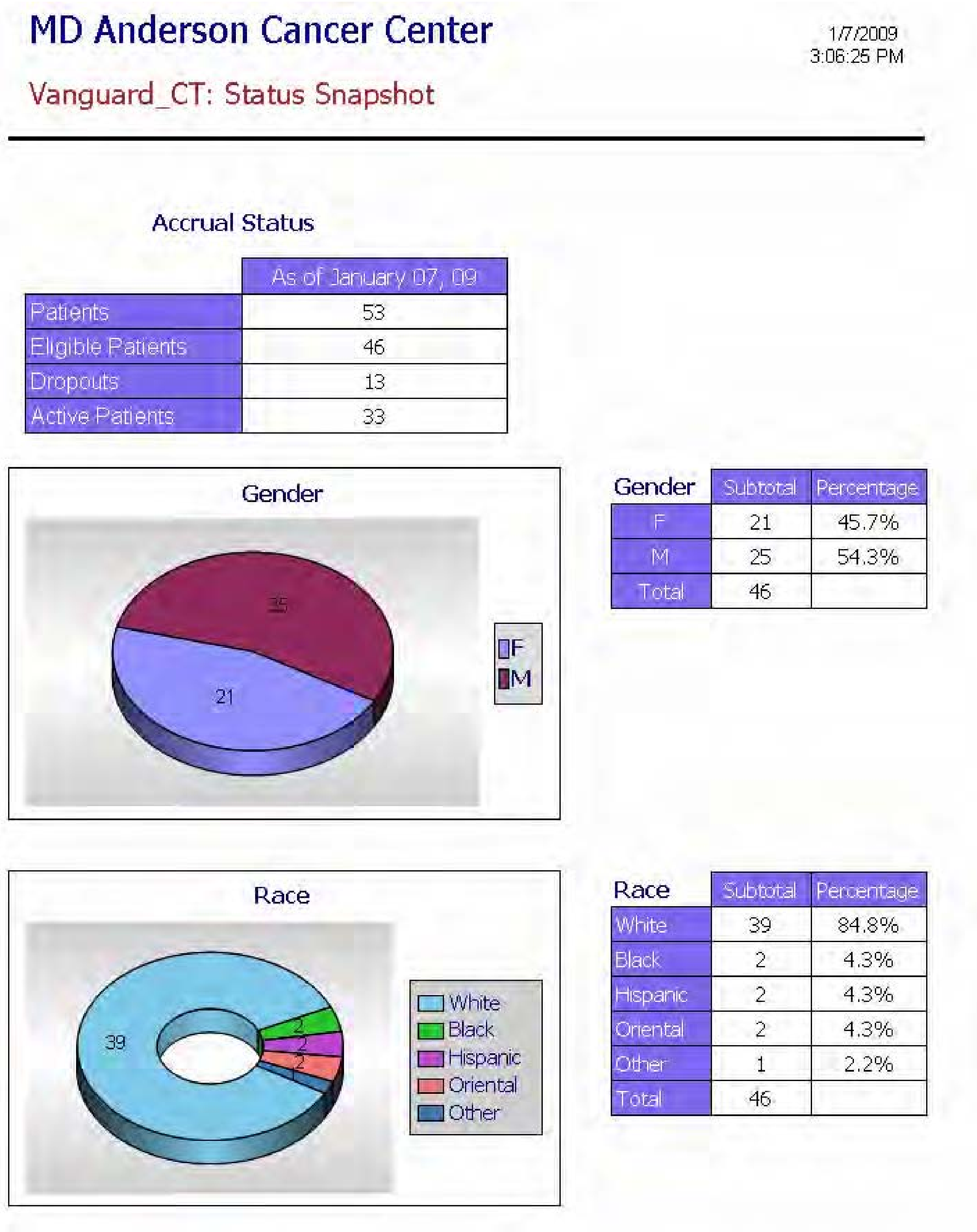
Page 1of 14

11) The example of the ReVITALization's Excel reports.



	K	L	M	N	O	P	Q
1	**Specimen Avail**	**Surgical Date**	**Single Wedge**	**Multiple Wedge**	**Single Segmentectomy**	**Multiple Segmentectomy**	**Lobectomy**
2							
3							
4							
5		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
6		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
7		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
8		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
9							
10							
11							
12							
13		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
14		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
15	FALSE						
16	TRUE						
17							
18							
19		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
20							
21							
22							
23							
24							
25							
26		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
27							
28							
29							
30							
31							
32							
33							
34							
35							

12) Vanguard Study: Status SnapShot



MD Anderson Cancer Center

1/7/2009
3:06:25 PM

Vanguard_CT: Status Snapshot

Accrual Status

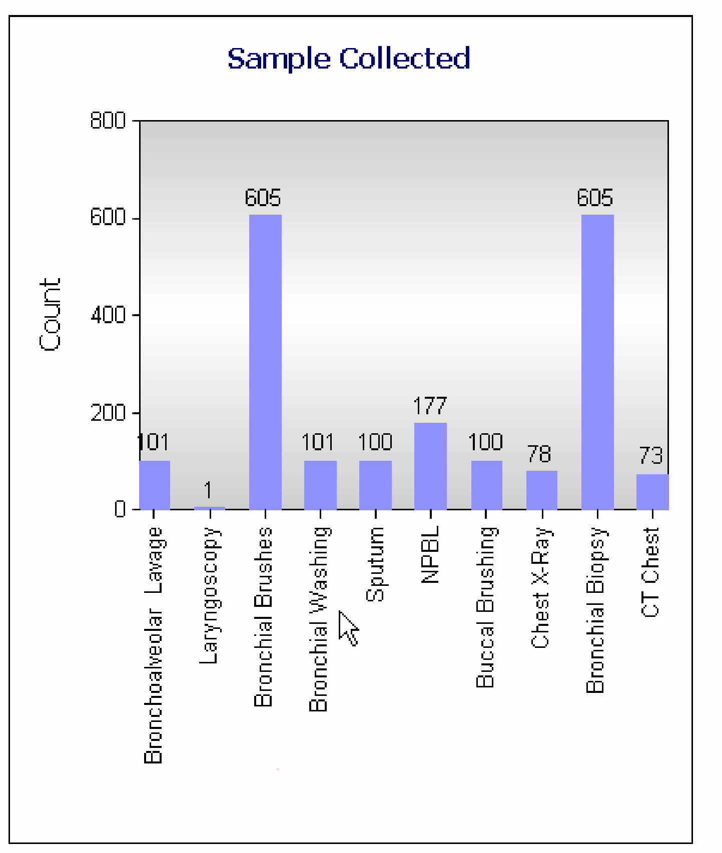
	As of January 07, 09
Patients	53
Eligible Patients	46
Dropouts	13
Active Patients	33

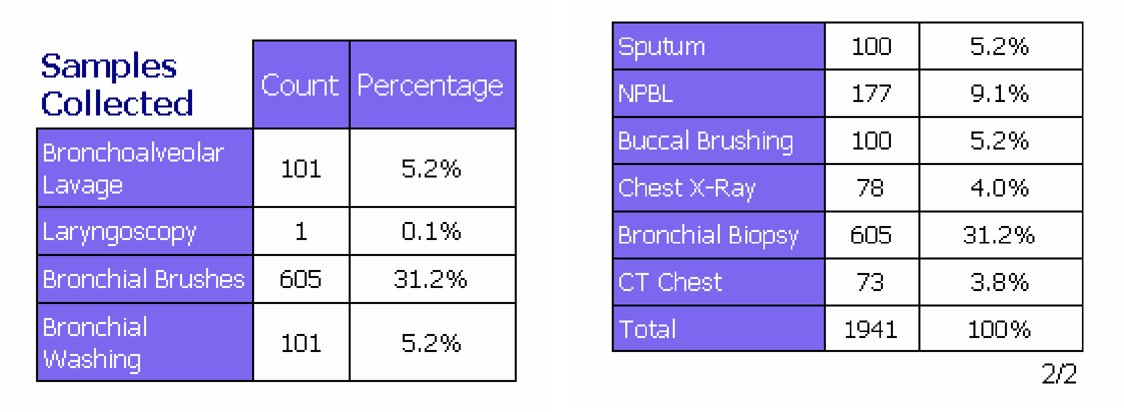
Gender	Subtotal	Percentage
F	21	45.7%
M	25	54.3%
Total	46	

Race	Subtotal	Percentage
White	39	84.8%
Black	2	4.3%
Hispanic	2	4.3%
Oriental	2	4.3%
Other	1	2.2%
Total	46	

Gender
25
21
F
M
Race
39
2
2
2
1
White
Black
Hispanic
Oriental
Other

13) Vanguard Study: Sample Collection

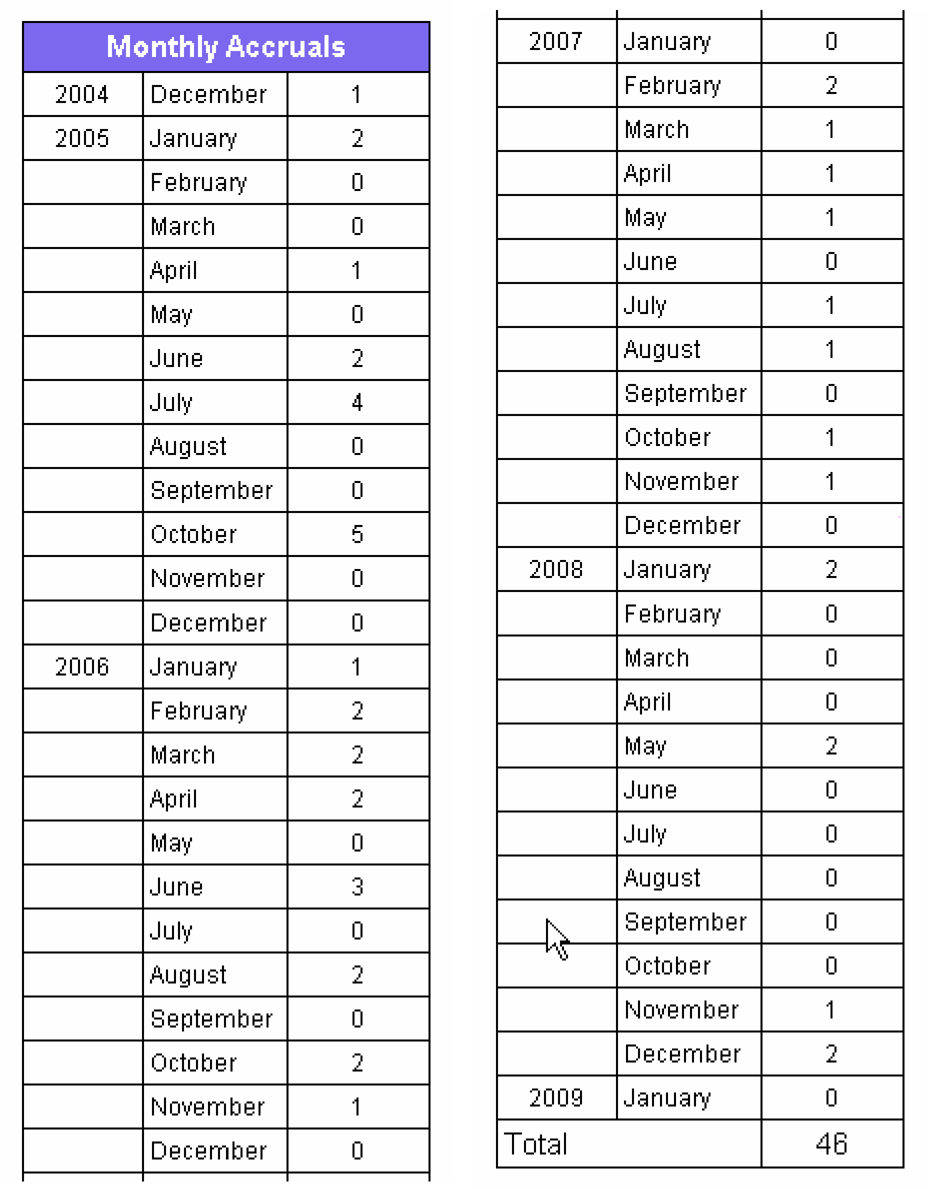




Samples Collected	Count	Percentage
Bronchoalveolar Lavage	101	5.2%
Laryngoscopy	1	0.1%
Bronchial Brushes	605	31.2%
Bronchial Washing	101	5.2%
Sputum	100	5.2%
NPBL	177	9.1%
Buccal Brushing	100	5.2%
Chest X-Ray	78	4.0%
Bronchial Biopsy	605	31.2%
CT Chest	73	3.8%
Total	1941	100%

2/2

14) Vanguard Study: Monthly Accrual



Monthly Accruals		
2004	December	1
2005	January	2
	February	0
	March	0
	April	1
	May	0
	June	2
	July	4
	August	0
	September	0
	October	5
	November	0
	December	0
2006	January	1
	February	2
	March	2
	April	2
	May	0
	June	3
	July	0
	August	2
	September	0
	October	2
	November	1
	December	0
2007	January	0
	February	2
	March	1
	April	1
	May	1
	June	0
	July	1
	August	1
	September	0
	October	1
	November	1
	December	0
2008	January	2
	February	0
	March	0
	April	0
	May	2
	June	0
	July	0
	August	0
	September	0
	October	0
	November	1
	December	2
2009	January	0
Total		46

Deregulated *EGFR* Signaling during Lung Cancer Progression: Mutations, Amplicons, and Autocrine Loops

Adi F. Gazdar and John D. Minna

One or more members of the family of epidermal growth factor receptor (*EGFR*) genes are overexpressed or otherwise deregulated in virtually all epithelial tumors, including non–small cell lung cancers (NSCLC). This and related observations on the importance of protein phosphorylation and the discovery that the first identified oncogene, v-Src, is a protein kinase led John Mendelsohn and Gordon Sato to select *EGFR* as the first target of molecular targeted therapy more than 20 years ago (1, 2). EGFR family members are deregulated in cancers by the following three fundamental mechanisms: activating gene mutations, increased gene copy number (via amplification or polysomy), and altered ligand expression (with possible formation of autocrine loops; ref. 3). Two reports in this issue of the journal advance our understanding of the role of all three mechanisms in the pathogenesis and progression of NSCLC (4, 5). Before discussing these reports, however, we will present background information on *EGFR* signaling and its deregulation in cancers.

Reversible protein phosphorylation as a crucial regulator of many essential cell functions has been elucidated over the past 50 years. A superfamily of more than 500 highly conserved protein kinase genes contains about 2% of the genome (6). Specific kinases phosphorylate serine/threonine or tyrosine residues or have dual specificity. The tyrosine kinases, which catalyze the transfer of γ phosphate of ATP to tyrosine residues on protein substrates, fall into two classes: transmembrane receptors (receptor tyrosine kinase) and nonreceptors. Subclass I of the receptor tyrosine kinases is the EGFR family, which consists of four members: *EGFR* (or *EGFR1*, *ERBB2*, *HER1*), *EGFR2* (or *ERBB2*, *HER2*), *EGFR3* (or *ERBB3*, *HER3*), and *EGFR4* (or *ERBB4*, *HER4*; ref. 3). Receptor-ligand interaction results in formation of homodimers or heterodimers (between family members), activation of the intrinsic kinase domain, and phosphorylation of specific tyrosine residues in the cytoplasmic tail of the receptor. The phosphorylated residues become docking sites for multiple proteins, which in turn activate downstream signaling pathways including the *PI3K/AKT* prosurvival, *STAT* transcription, and *RAS/RAF/MEK* proliferation pathways.

Eleven members of the EGF family have been identified as ligands for the EGFR family. *HER2* is not ligand activated because of its unique extracellular spatial structure but is the preferred dimerization partner for other family members; its heterodimers preferentially enhance ligand binding (7). EGFR3 is "kinase dead" (i.e., it lacks intrinsic kinase activity) and, as with HER2, functions via heterodimerization. The EGF ligands show specificity for multiple homodimers or heterodimers (7). Epiregulin is a pan-EGFR family ligand that preferentially activates heterodimeric receptor complexes (8). The EGF ligands are produced as transmembrane precursors that are cleaved into their soluble forms by proteases ("sheddases") of the ADAM family (especially ADAM10 and ADAM17) or by matrix metalloproteinases, a process known as ectodomain shedding (9). Other receptor pathways also may activate EGFR signaling by activating the EGFR pathway via "cross talk" and/or "transactivation." An important new example of this with relevance to EGFR is the inflammatory cytokine interleukin-6, which activates the Janus-activated kinase/signal transducer and activator of transcription system, which in turn activates EGFR pathway signaling. High levels of interleukin-6 have been described in many cancers, including EGFR-mutant lung cancers, providing an additional method for EGFR activation and a new therapeutic target.

NSCLC cells can produce and release several of the EGF ligands (10–12). Under certain circumstances, the membrane-anchored isoforms and soluble growth factors also may act as biologically active ligands. Therefore, depending on the circumstances, these ligands may induce juxtacrine, autocrine, paracrine, and/or endocrine signaling (13). Establishing EGFR autocrine loops renders the cells sensitive to inhibition by tyrosine kinase inhibitors (10, 12). Zhou et al. (14) described the presence of an autocrine heregulin-EGFR3 loop associated with up-regulation of the sheddase ADAM10. Inhibiting ADAM10 with a specific inhibitor prevented the processing and activation of multiple EGF ligands. Recent reports indicate that breast and NSCLC cells (especially those with EGFR mutations) may produce large amounts of interleukin-6, activating another autocrine loop that drives tumorigenesis (15, 16).

Mutations of *EGFR* may target many regions of the gene, especially the extracellular domain in glioblastomas (17) and the kinase domain in lung cancers (18, 19). *EGFR* mutations may play a major role in lung tumorigenesis but also leave lung tumor cells dependent on EGFR signaling pathway activation for growth and survival ("oncogene addiction"; refs. 19, 20). Therefore, inhibition of *EGFR* signaling by tyrosine kinase inhibitors rapidly leads to apoptosis and growth cessation. In the 4 years since the discovery of the mutations, however, it was realized that primary tumor response and resistance to tyrosine kinase inhibitors are influenced by many factors, including mutations, mutation type, and copy numbers of *EGFR*; *EGFR3* activation; *KRAS* mutations; *MET* amplification, and others (21–23). Therefore, although some studies

---

Authors' Affiliations: Hamon Center for Therapeutic Oncology Research and the Simmons Cancer Center, UT Southwestern Medical Center, Dallas, Texas

Received 04/15/2008; accepted 04/22/2008.

Grant support: NCI Lung Cancer Specialized Program of Research Excellence grant P50CA70907, Early Detection Research Network, National Cancer Institute, and Department of Defense VITAL and PROSPECT grants.

Requests for reprints: Adi F. Gazdar, UT Southwestern Medical Center, Building NB8.206, 6000 Harry Hines Boulevard, Dallas, TX 75390-8593. Phone: 214-648-4921; Fax: 214-648-4940; E-mail: adi.gazdar@utsouthwestern.edu.

©2008 American Association for Cancer Research.

doi:10.1158/1940-6207.CAPR-08-0080

(usually from single institutions analyzing highly selected patient populations) have shown very high response rates of EGFR-mutant tumors to tyrosine kinase inhibitors, large multi-institutional clinical trials have often failed to show a survival benefit of this approach, although increased copy number of *EGFR* (and *HER2* in some series) was associated with a good treatment outcome (24, 25). Although *EGFR* mutations and copy number gains may occur independently, they occur together more frequently than alone (26, 27). In addition, as with glioblastomas (17), the mutant allele is preferentially amplified in such cases (26). Therefore, "triple whammy" tumors (i.e., those with mutations, copy number gains, and mutant allele-specific amplifications) are in all probability highly oncogene addicted and likely to show dramatic and sustained responses to appropriate targeted therapies. Autocrine loops and other derangements of EGFR signaling are frequent in all forms of NSCLC, which therefore may involve tumors with more than three EGFR aberrations, or "multiple whammy" tumors.

The finding that all of these different mechanisms activate EGFR signaling in lung cancers signifies the presence and great importance of strong selective pressures on the EGFR signaling pathway in these cancers. This selectivity was dramatically highlighted by the finding of *EGFR* tyrosine kinase domain mutations, but lung cancer use of all these alternative mechanisms is equally important in underscoring the key role of the EGFR pathway in driving lung cancer pathogenesis. Of course, these findings also highlight how versatile tumor cells are in finding ways to activate the pathway. On a related note, the relapse and subsequent drug resistance of lung cancers that had responded to EGFR-targeting drugs (such as EGFR tyrosine kinase inhibitors) show the resourcefulness of these cancers in finding other ways to use the EGFR or other pathways (e.g., *KRAS*, c-MET) to ward off extinction. Relapse and resistance also highlight the need for tools that can determine whether the pathway is active in and identify "sensitive" therapeutic target(s) for individual lung cancers. It is also important to realize that the target is constantly changing, and thus different therapeutic options are needed at different disease stages.

We now evaluate the contributions of the articles by Zhang et al. (4) and Tang et al. (5) in the context of the EGFR signaling background detailed above. Lung cancer has a high mortality that usually is due to the development of metastatic lesions. Although relatively few studies have directly compared the molecular changes in primary tumors with those in corresponding metastatic tumors, the metastatic phenotype is characterized by changes in multiple cellular pathways (28). The study by Zhang et al. (4) was stimulated by previous work from their laboratory showing that epiregulin is one of the several highly expressed EGF ligands in EGFR-mutant NSCLC cells (10). This group tested the hypothesis that epiregulin is involved in the development of the metastatic phenotype. Immunostaining studies confirmed their previous observation that primary NSCLC tumors with localized disease stages frequently (in 65% of cases) expressed the ligand. They reported a significant correlation between ligand expression and advanced nodal stage (stage II) and a trend toward shorter survival. *In vitro* studies confirmed the role of epiregulin in promoting tumor growth and invasion. These analyses show a clear role for epiregulin in tumor cell survival, invasion, and metastasis. Because epiregulin can stimulate multiple members of the *EGFR* receptor family, activation of both *EGFR* and *EGFR3* signaling may contribute to carcinogenesis. Because ligand expression is much more frequent than are *EGFR* mutations or copy number gains, these findings provide further evidence that autocrine loops may be the major mechanism by which *EGRF* signaling is deregulated in all histologic forms of NSCLC. Future studies should comprehensively analyze all 11 EGF ligands found in lung cancers because other members of this ligand group may have similar tumor-promoting actions.

As mentioned earlier, *EGFR* mutations and copy gains occur frequently in the same tumors. Previous studies have shown widespread field effects throughout the respiratory epithelium of smokers (29, 30), suggesting that tobacco exposure damages the entire respiratory epithelium. Most *EGFR* mutations occur in lung cancers of lifetime never smokers, which have a largely unknown etiology (31). In their earlier work, the authors carefully microdissected histologically normal respiratory epithelium from small airways surrounding mutation-containing tumors (32); often present in airways within or near the tumor but seldom in distant sites, the mutations reflected a limited field effect. Therefore, exposure and damage seem to be much more limited in never smokers than in current or former smokers. In their present study, Tang et al. conducted a more extensive field study, assessing the presence of mutations and copy number gains (by fluorescence *in situ* hybridization technique) in primary NSCLC, corresponding metastases, and histologically normal respiratory epithelium. As in their previous study, mutations and EGFR protein overexpression were a localized field effect. The key present findings are that copy number gains were absent in normal epithelium and were distributed heterogeneously in primary tumors and more evenly in metastases. Tang et al. (5) have answered the question, "Which came first, the chicken (copy number gains) or the egg (mutations)?" The finding of mutant allele-specific gains gives the nod to the egg.

The prototype *EGFR* gene is not the only EGFR pathway gene amplified in NSCLC. A recent report describes amplification of other pathway members including *HER2*, *SHC1*, and *AKT* (33). Our unpublished work indicates that other pathway genes including *KRAS* and *BRAF* may also be amplified in NSCLC. Although mutations of pathway genes are usually mutually exclusive, single tumors may contain copy number gains for multiple genes or a single pathway mutation and one or more pathway gene copy number gains.[1]

Two other recently published studies (34, 35) are consistent with the findings of Tang et al. (5). Cancers arise as a result of multistage processes, and a lesion known as atypical adenomatous hyperplasia is recognized as a precursor or premalignant lesion for peripheral lung adenocarcinomas. Atypical adenomatous hyperplasia lesions progress to noninvasive cancers known as bronchioloalveolar carcinomas as defined by the strict criteria of WHO classification (36). Bronchioloalveolar carcinoma tumors may become invasive and eventually metastatic. Early invasive cancers may contain invasive and noninvasive components that can be microdissected and examined separately. By examining the various stages of lung pathogenesis for *EGFR* mutations and copy gains, both reports (34, 35) conclude that mutations are early, preinvasive changes, whereas copy number gains are later events associated with the invasive phenotype (Fig. 1).

---

[1] Unpublished data.

All of these findings are consistent with the hypothesis that mutations precede copy number gains, which may be associated with the metastatic phenotype. Therefore, mutations are likely to show little or no heterogeneity in primary or metastatic tumors, and copy number gains may be absent or heterogeneously distributed in primary tumors and relatively evenly distributed within metastatic sites. Further work will be needed to confirm that copy number gains are part of the metastatic phenotype.

What are the clinical implications of these findings? The data of Zhang et al. (4) suggest that about two thirds of all NSCLCs express at least one of the EGF ligands. Testing the expression of the other 10 known ligands in this cohort presumably would have shown an even higher percentage. The

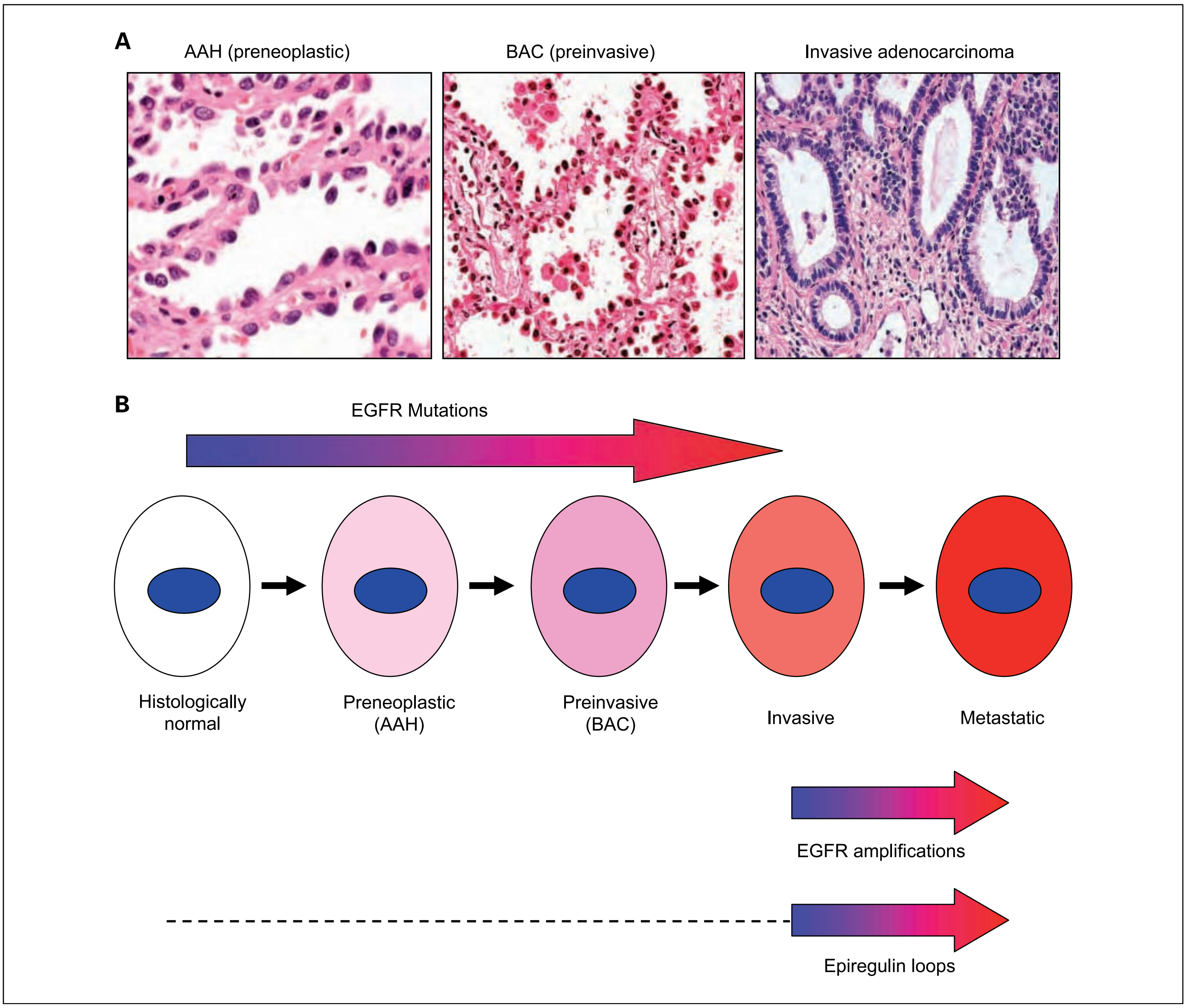


Fig. 1. Deregulation of the *EGFR* gene during the multistage pathogenesis of peripheral lung adenocarcinomas. *A,* peripheral adenocarcinomas are believed to arise from preneoplastic lesions known as atypical adenomatous hyperplasias (*AAH*), which first progress to a preinvasive neoplastic stage called bronchioloalveolar carcinoma (*BAC*). Foci of invasion may develop in the fibrotic centers of bronchioloalveolar carcinomas, which then are called invasive adenocarcinomas, although noninvasive elements may persist at the edges of the tumors. Metastases ultimately develop (not shown). *B,* from the article by Tang et al. (5) and from the literature cited in the text, *EGFR* mutations commence early during pathogenesis and can be detected in histologically normal respiratory epithelium near tumors (localized field effect). Mutations are more frequent in preneoplastic (atypical adenomatous hyperplasia) and preinvasive (bronchioloalveolar carcinoma) stages than in normal epithelium. Therefore, there is relatively little heterogeneity of mutations in invasive carcinomas, and the mutations contribute to tumor pathogenesis. In contrast, gene copy number gains, often in the form of amplifications, commence relatively late in pathogenesis, usually at the tumor stage. They are more frequent in metastatic lesions, suggesting that they may be progression events involved in the metastatic phenotype. Much less is known about the timing of epiregulin loops (either autocrine, paracrine, or juxtacrine). From the data of Zhang et al. (4), however, it would seem that epiregulin loops can be detected in primary invasive tumors but are more frequent or active during the metastatic stage. The dashed line indicates that the timing of the appearance of these loops during earlier preinvasive stages is unknown.

expression of EGFR protein in most NSCLCs, including squamous cell carcinomas, raises the question of what mechanism causes deregulation. Mutations and copy number gains explain only a minority of these cases and probably are not important mechanisms in squamous cell carcinomas. As suggested by the data of Zhang et al. (4), activation of autocrine (or paracrine or juxtacrine) loops is an attractive alternative mechanism. If this loop is dependent on continued EGFR signaling and is inhibited by tyrosine kinase inhibitor therapy, as suggested by the data, this would be a plausible explanation for why some nonmutant tumors of all histologic types with nearly diploid copy number respond to tyrosine kinase inhibitor therapy (24, 37). Future retrospective and prospective studies are needed to determine whether EGF ligand expression is an additional predictive factor for tyrosine kinase inhibitor response. The concept that the driving force behind many or most NSCLC tumors is EGF ligand receptor loops offers the clinician the following additional avenues for potential targeted therapies: preventing sheddase up-regulation or activity, preventing ligand production directly or by inhibition of the loop at a more upstream stage, targeting the soluble form of the ligand, and preventing ligand-receptor interaction.

With the identification of deregulated expression of EGF family ligands in lung cancer pathogenesis, we can now consider using the relevant ligands for early cancer diagnosis, identifying key therapeutic targets, and as biomarkers to monitor response to chemoprevention or very early treatment. Because the ligands are soluble, they potentially could be detected in blood or bronchial lavage specimens in addition to biopsy and brushing specimens. Furthermore, while exploring their diagnostic and therapeutic targeting roles, we need to understand the molecular mechanisms leading to the deregulated expression of these ligands. Copy number changes, mutations, promoter alterations (including epigenetic changes), the role of specific transcription factors (such as the lineage-specific oncogene *TITF1*), and altered miRNA expression are all potential mechanisms that need to be explored, as does ligand expression in cancer stem cells.

Another major clinical interest is to understand the sequential appearance of molecular changes during multistage pathogenesis. The appearance of *EGFR* mutations at a preinvasive and even at a premalignant phase creates opportunities to use *EGFR* mutation markers for risk identification, early detection, and prevention, particularly for never smokers, who are at most risk for *EGFR*-mutant tumors and for whom no such markers currently exist (31). Early *EGFR* mutations also have important implications for the study of EGFR tyrosine kinase inhibitors in the adjuvant/second primary tumor prevention setting. Whereas mutations seem to be initiating events, copy number gains are related to progression and metastatic events. Therefore, heterogeneity may occur both within the primary tumor and between the primary tumor and metastatic sites. These considerations are important if copy number gains are used as a marker for selecting targeted therapies, and they indicate the importance of testing for this marker in tumor samples obtained immediately before therapy versus relying on marker data from earlier samples.

The reports of Zhang et al. and Tang et al. in this issue of the journal shed new light on the highly complex, multifaceted, and as yet incompletely understood nature of the *EGFR* signaling pathway. This pathway in NSCLCs and in the bronchial epithelium of patients at a high lung-cancer risk will be a critical focus of diagnostic, preventive, and therapeutic efforts for the foreseeable future.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Cohen P. Protein kinases-the major drug targets of the twenty-first century? Nat Rev Drug Discov 2002;1:309–15.

2. Mendelsohn J. Antibody-mediated EGF receptor blockade as an anticancer therapy: from the laboratory to the clinic. Cancer Immunol Immunother 2003;52:342–6.

3. Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. Nature reviews 2005;5:341–54.

4. Zhang J, Iwanaga K, Choi CC, et al. Intratumoral epirgulin is a marker of advanced disease in non-small cell lung cancer patients and confers invasive properties on EGFR-mutant cells. Cancer Prev Res. In press 2008.

5. Tang X, Varella-Garcia M, Xavier AC, et al. EGFR abnormalities in the pathogenesis and progression of lung adenocarcinoma. Cancer Prev Res. In press 2008.

6. Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci 2002;27:514–20.

7. Citri A, Yarden Y. EGF-ERBB signalling: towards the systems level. Nat Rev Mol Cell Biol 2006;7:505–16.

8. Shelly M, Pinkas-Kramarski R, Guarino BC, et al. Epiregulin is a potent pan-ErbB ligand that preferentially activates heterodimeric receptor complexes. J Biol Chem 1998;273:10496–505.

9. Hynes NE, Schlange T. Targeting ADAMS and ERBBs in lung cancer. Cancer cell 2006;10:7–11.

10. Fujimoto N, Wislez M, Zhang J, et al. High expression of ErbB family members and their ligands in lung adenocarcinomas that are sensitive to inhibition of epidermal growth factor receptor. Cancer Res 2005;65:11478–85.

11. Volante M, Saviozzi S, Rapa I, et al. Epidermal growth factor ligand/receptor loop and downstream signaling activation pattern in completely resected nonsmall cell lung cancer. Cancer 2007;110:1321–8.

12. Wu W, O'Reilly MS, Langley RR, et al. Expression of epidermal growth factor (EGF)/transforming growth factor-α by human lung cancer cells determines their response to EGF receptor tyrosine kinase inhibition in the lungs of mice. Mol Cancer Ther 2007;6:2652–63.

13. Singh AB, Harris RC. Autocrine, paracrine and juxtacrine signaling by EGFR ligands. Cell Signal 2005;17:1183–93.

14. Zhou BB, Peyton M, He B, et al. Targeting ADAM-mediated ligand cleavage to inhibit HER3 and EGFR pathways in non-small cell lung cancer. Cancer Cell 2006;10:39–50.

15. Schafer ZT, Brugge JS. IL-6 involvement in epithelial cancers. J Clin Invest 2007;117:3660–3.

16. Gao SP, Mark KG, Leslie K, et al. Mutations in the EGFR kinase domain mediate STAT3 activation via IL-6 production in human lung adenocarcinomas. J Clin Invest 2007;117:3846–56.

17. Frederick L, Eley G, Wang XY, James CD. Analysis of genomic rearrangements associated with EGRFvIII expression suggests involvement of Alu repeat elements. Neuro-oncol 2000;2:159–63.

18. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006;118:257–62.

19. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nature reviews 2007;7:169–81.

20. Gazdar AF, Shigematsu H, Herz J, Minna JD. Mutations and addiction to EGFR: the Achilles 'heal' of lung cancers? Trends Mol Med 2004;10:481–6.

21. Engelman JA, Cantley LC. The role of the ErbB family members in non-small cell lung cancers sensitive to epidermal growth factor receptor kinase inhibitors. Clin Cancer Res 2006;12:4372–6s.

22. Pao W, Wang TY, Riely GJ, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS Med 2005;2:e17.

23. Riely GJ, Politi KA, Miller VA, Pao W. Update on epidermal growth factor receptor mutations in non-small cell lung cancer. Clin Cancer Res 2006;12:7232–41.

24. Shepherd FA, Rosell R. Weighing tumor biology in treatment decisions for patients with non-small cell lung cancer. J Thorac Oncol 2007;2 Suppl 2: S68–76.

25. Cappuzzo F, Varella-Garcia M, Shigematsu H, et al. Increased HER2 gene copy number is associated with response to gefitinib therapy in epidermal growth factor receptor-positive non-small-cell lung cancer patients. J Clin Oncol 2005; 23:5007–18.

26. Nomura M, Shigematsu H, Li L, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007;4:e125.

27. Okabe T, Okamoto I, Tamura K, et al. Differential constitutive activation of the epidermal growth factor receptor in non-small cell lung cancer cells bearing EGFR gene mutation and amplification. Cancer Res 2007;67:2046–53.

28. Minn AJ, Gupta GP, Siegel PM, et al. Genes that mediate breast cancer metastasis to lung. Nature 2005;436:518–24.

29. Mao L, Lee JS, Kurie JM, et al. Clonal genetic alterations in the lungs of current and former smokers [see comments]. J Natl Cancer Inst 1997;89: 857–62.

30. Wistuba II, Lam S, Behrens C, et al. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89:1366–73.

31. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. Nat. Rev 2007;7:778–90.

32. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65:7568–72.

33. Lockwood WW, Chari R, Coe BP, et al. DNA amplification is a ubiquitous mechanism of oncogene activation in lung and other cancers. Oncogene. In press 2008.

34. Soh J, Toyooka S, Ichihara S, et al. Sequential molecular changes during multistage pathogenesis of small peripheral adenocarcinomas of the lung. J Thorac Oncol 2008;3:340–7.

35. Yatabe Y, Takahashi T, Mitsudomi T. Epidermal growth factor receptor gene amplification is acquired in association with tumor progression of EGFR-mutated lung cancer. Cancer Res 2008;68: 2106–11.

36. Brambilla E, Travis WD, Colby TV, Corrin B, Shimosato Y. The new World Health Organization classification of lung tumours. Eur Respir J 2001; 18:1059–68.

37. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 2005;353:133–44.

Confidence Intervals of Interaction Index for Assessing Multiple Drug Interaction

J. JACK LEE and MAIYING KONG

Studies of interactions among biologically active agents have become increasingly important in many branches of biomedical research. We consider that the Loewe additivity model is one of the best general reference models for defining drug interactions. Based on the Loewe additivity model, synergy occurs when the interaction index is less than one, and antagonism occurs when interaction index is greater than one. Starting from the Loewe additivity model and the marginal dose–effect curve for each drug involved in a combination, we first present a procedure to estimate the interaction index and its associated confidence interval at a combination dose with observed effects. Following Chou and Talalay's method for assessing drug interaction based on the plot of interaction indices versus effects for combination doses at a fixed ray, we then construct a pointwise $(1-\alpha)\times 100\%$ confidence bound for the curve of interaction indices versus effects. We found that these methods work better on the logarithm transformed scale than on the untransformed scale of the interaction index. We provide simulations and case studies to illustrate the performances of these two procedures, and present their pros and cons. We also provide S-Plus/R code to facilitate the implementation of these two procedures.

Key Words: Antagonism; Loewe additivity model; Synergy.

1. Introduction

Studies of interactions among biologically active agents, such as drugs, carcinogens, or environmental pollutants, have become increasingly important in many branches of biomedical research. Our research group reviewed the literature (Lee, Kong, Ayers, and Lotan 2007) and agree with many researchers (e.g., Berenbaum 1985, 1989; Greco, Bravo, and Parsons 1995) that the Loewe additivity model should be considered as the "gold standard" for defining drug interactions

For a combination of k drugs ($k \geq 2$) at $(d_1, \ldots, d_k)$, based on the Loewe additivity model, drug interactions at this combination can be characterized as

$$\frac{d_1}{D_{y,1}} + \cdots + \frac{d_k}{D_{y,k}} \begin{cases} < 1, & \text{synergy;} \\ = 1, & \text{additivity;} \\ > 1, & \text{antagonism.} \end{cases} \qquad (1)$$

Here $d_1, \ldots, d_k$ are doses of each drug in the mixture of the k drugs resulting in effect y, and $D_{y,1}, \ldots, D_{y,k}$ are the doses of drugs that result in the effect y for each respective drug given alone. The summation, $\frac{d_1}{D_{y,1}} + \cdots + \frac{d_k}{D_{y,k}}$, is called the interaction index, which is denoted as τ. Based on the Loewe additivity model, the combination dose $(d_1, \ldots, d_k)$ is said to be synergistic if the interaction index is less than the constant number of 1, and additive or antagonistic if the index is equal to or greater than 1, respectively. To give an intuitive idea about the interaction index, we illustrate its meaning in the special case of $k = 2$. Note that the combination dose (d_1, d_2) produces the same effect y as drug 1 alone at dose level $D_{y,1}$, and drug 2 alone at dose level $D_{y,2}$, which implies that 1 unit of drug 2 will produce the same effect as $\frac{D_{y,1}}{D_{y,2}}$ units of drug 1. Thus, the amount of dose at the combination (d_1, d_2) equals to $d_1 + d_2\frac{D_{y,1}}{D_{y,2}}$

© American Statistical Association
Statistics in Biopharmaceutical Research
February 2009, Vol. 1, No. 1
DOI: 10.1198/sbr.2009.0001

in terms of drug 1 dose. By definition, $\tau = \frac{d_1}{D_{y,1}} + \frac{d_2}{D_{y,2}}$, which implies that $d_1 + d_2\frac{D_{y,1}}{D_{y,2}} = \tau D_{y,1}$. $\tau < 1$ implies $d_1 + d_2\frac{D_{y,1}}{D_{y,2}} = \tau D_{y,1} < D_{y,1}$, therefore, the amount of the combination dose (d_1, d_2) to produce the same effect y is less than the amount of dose when single drug is applied, hence, indicating synergy. The smaller τ is, the more is the reduction of the amount of dose in the combination, and the stronger is the synergy. Similarly, $\tau > 1$ implies that $d_1 + d_2\frac{D_{y,1}}{D_{y,2}} = \tau D_{y,1} > D_{y,1}$, that is, the amount of the combination dose (d_1, d_2) producing the same effect y is more than each single drug dose, hence indicating antagonism. The geometric interpretation of the interaction index can be best shown graphically in Figure 1, panels (A) and (B), where $P = (D_{y,1}, 0)$, $Q = (0, D_{y,2})$, $U = (d_1, d_2)$, all yield the same effect y. If we draw a line $\overline{RS}$ passing through U and parallel to the line $\overline{PQ}$, and draw a line $\overline{OU}$ intercepting with $\overline{PQ}$ at V, then, from the basic geometric properties, the interaction index can be expressed as the ratio of $\frac{\text{length}(\overline{OU})}{\text{length}(\overline{OV})}$. In Figure 1(A), the closer the point U is toward the origin, the less is the amount of combination dose required to produce the same effect as drug 1 alone at $D_{y,1}$ or drug 2 alone at $D_{y,2}$, hence, the stronger is the synergy. By the same token in Figure 1(B), the further the point U is away from the origin, the larger is the amount of combination dose required to produce the same effect as drug 1 alone at $D_{y,1}$ or drug 2 alone at $D_{y,2}$, hence, the stronger is the antagonism. From Figure 1, we conclude that the interaction index can be used to measure the mode and magnitude of drug interactions.

Given the combination dose $(d_1, \ldots, d_k)$ and its effect y, and the marginal dose–effect curves $f_i(D_i)$ for drug i $(i = 1, \ldots, k)$, the calculation of the interaction index at a combination dose $(d_1, \ldots, d_k)$ is straightforward. One simply replaces $D_{y,i}$ by $f_i^{-1}(y)$, where f_i^{-1} is the inverse function of f_i $(i = 1, \ldots, k)$. However, since the dose–effect curves are usually estimated and the effect y is observed with error, to make valid inferences for drug interactions, one needs to account for all these variabilities. In other words, one needs to consider the estimated interaction indices along with their variances to make valid statistical inferences on drug interaction.

In most settings, the functional form of the marginal dose response curves may not be known and need to be estimated from the data. Chou and Talalay (1984) proposed the median-effect equation which has been widely used to model the dose–effect curve with good success (Chou 2006). In our cell line study (Kong and Lee 2006; Lee et al. 2007), we also find that the median-effect equation fits the data well. Chou and Talalay's median-effect equation has the following form

$$E = \frac{\left(\frac{d}{D_m}\right)^m}{1 + \left(\frac{d}{D_m}\right)^m}, \tag{2}$$

where d is the dose of a drug eliciting effect E, D_m is the median effective dose of a drug, and m is a slope parameter depicting the shape of the curve. When m is negative, the curve described by Equation (2) falls with increasing drug concentration; when m is positive, the curve rises with increasing drug concentration. The median-effect Equation (2) is independent of the drug's mechanisms of action and does not require knowledge of conventional kinetic constants (Chou 2006; Greco et al. 1995). Under the assumption that the dose–effect curves follow Chou and Talalay's median-effect equation, in Section 3 we investigate the characteristics of the interaction index and its logarithmic transformation, and propose a procedure to construct the confidence interval for the estimated interaction index.

Although interaction index can be estimated at each combination dose separately, this approach is not efficient. The result tends to be more varying as it depends on only measurement at a single combination dose level. To gain efficiency, one can assume a model and pool data at various combination doses to form a better estimate of the interaction index. One commonly used approach is applying the ray design. Chou and Talalay (1984) and Chou (1991) proposed a procedure to characterize a two-drug interaction by first fitting marginal dose–effect curves and a dose–effect curve for the combination doses with their components at a fixed ray (i.e., d_1/d_2=a constant, forming a ray in the $d_1 \times d_2$ dose plane), then assessing drug interaction based on the plot of their combination indices versus effects for combination doses at this fixed ray. The confidence intervals for the combination indices were constructed by Monte Carlo techniques (*CalcuSyn* at *http://www.biosoft.com/w/calcusyn.htm*; Belen'kii and Schinazi 1994; *CompuSyn* at *http://www.combosyn.com/*). The combination index has the same form as the interaction index when the combined drugs are mutually exclusive. However, Chou and Talalay's mutual exclusiveness and nonexclusiveness criteria are difficult to evaluate, and the combination index has been criticized by many researchers (Berenbaum 1989; Greco et al., 1995). In Section 3, by adopting the interaction index for Chou and Talalay's method, we extend their method to assess drug interactions among $k(\geq 2)$ drugs, and propose a procedure to construct a pointwise $(1 - \alpha) \times 100\%$ confidence bound for the

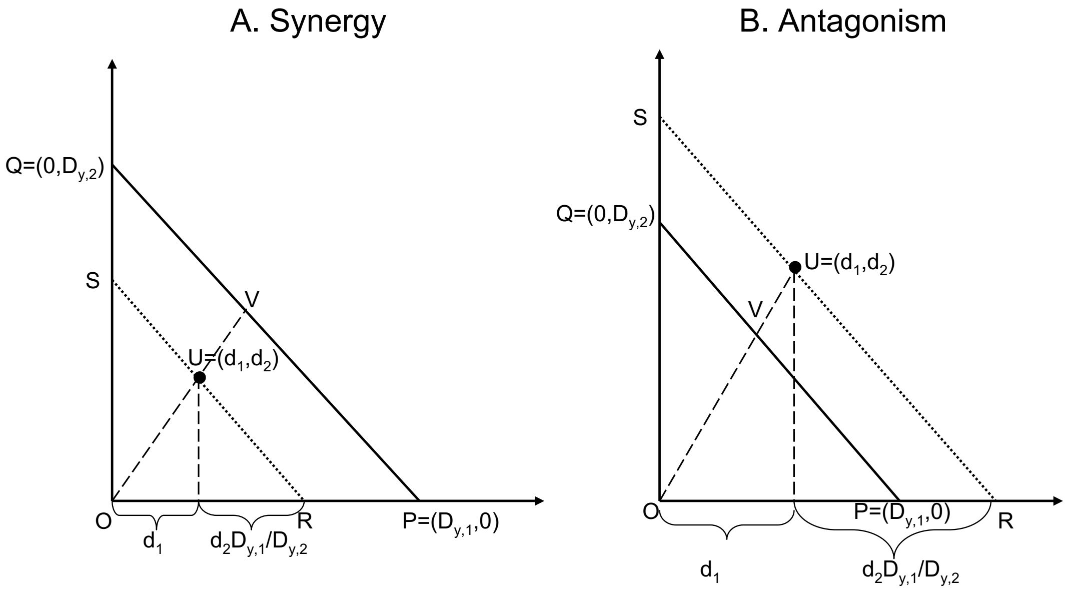


Figure 1. Illustration of the interaction index. $P = (D_{y,1}, 0)$ represents the drug 1 dose producing effect y, $Q = (0, D_{y,2})$ represents the drug 2 dose producing effect y, and $U = (d_1, d_2)$ represents the combination dose producing the same effect y. $\overline{RS}$ is the line passing by U and parallel to $\overline{PQ}$, and V is the intercept of $\overline{OU}$ and $\overline{PQ}$. The interaction index can be expressed as the ratio of $\frac{\text{length}(\overline{OU})}{\text{length}(\overline{OV})}$. Panel A illustrates the case for synergy while Panel B for antagonism.

estimated curve of interaction indices versus effects by accounting for all the variabilities in estimating the dose–effect curves. In Section 4, we present the results from simulations and case studies to show that our proposed procedure performs at least as well as the Monte Carlo procedure in terms of covering the underlying curves and shortening the confidence bound, and performs better in terms of taking less time to compute. The last section is devoted to discussion.

2. Interaction Index and its Confidence Interval at a Combination Dose

In this section, we will present how to construct a confidence interval for interaction index at a combination dose $(d_1, \ldots, d_k)$ with an effect y. Here the dose–effect curve for drug i $(i = 1, \ldots, k)$ is estimated from the marginal data with only ith drug being applied. Note that Chou and Talalay's median-effect equation (2) also can be rewritten as

$$\log \frac{E}{1-E} = m(\log d - \log D_m) = \beta_0 + \beta_1 \log d, \quad (3)$$

where $\beta_0 = -m \log D_m$ and $\beta_1 = m$. The dose producing effect E can be written as either

$$d = D_m \left(\frac{E}{1-E} \right)^{\frac{1}{m}}, \quad (4)$$

or

$$d = \exp\left(-\frac{\beta_0}{\beta_1}\right) \left(\frac{E}{1-E} \right)^{\frac{1}{\beta_1}}. \quad (5)$$

Suppose model (3) has the form $\log \frac{E}{1-E} = \beta_0 + \beta_1 \log d + \epsilon$ with ϵ following $N(0, \sigma^2)$, then we may regress $\log \frac{E}{1-E}$ on $\log d$ to get the marginal dose–effect curve $\log \frac{E}{1-E} = \hat{\beta}_{0,i} + \hat{\beta}_{1,i} \log d$ for drug i with $i = 1, \ldots, k$. Meanwhile we may get the variances and covariances for the estimates $\hat{\beta}_{0,i}$ and $\hat{\beta}_{1,i}$ for $i = 1, \ldots, k$. If the observed mean effect at a combination dose $(d_1, \ldots, d_k)$ is y, then, based on (5), the associated interaction index can be estimated by

$$\hat{\tau} = \sum_{i=1}^{k} \frac{d_i}{\hat{D}_{y,i}} = \sum_{i=1}^{k} \frac{d_i}{\exp(-\frac{\hat{\beta}_{0,i}}{\hat{\beta}_{1,i}}) \left(\frac{y}{1-y}\right)^{\frac{1}{\hat{\beta}_{1,i}}}}. \quad (6)$$

The simulations in Section 4 indicate that the distribution of $\log(\hat{\tau})$ is approximately normal, while $\hat{\tau}$ deviates from a normal distribution for large σ's. Thus, we should apply the delta method (Bickel and Doksum 2001) to $\log \tau$ instead of τ, then we take the exponential transformation to get the confidence interval for τ. By applying delta method to $\log(\hat{\tau})$, we get

$$\begin{aligned}
&\operatorname{var}(\log(\hat{\tau}))\\
&\quad\simeq \frac{1}{\hat{\tau}^2}\operatorname{var}(\hat{\tau})\\
&\quad\simeq \frac{1}{\hat{\tau}^2}\left(\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{0,1}},\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{1,1}},\ldots,\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{0,k}},\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{1,k}},\frac{\partial\hat{\tau}}{\partial y}\right)\\
&\qquad\times\Sigma\left(\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{0,1}},\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{1,1}},\ldots,\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{0,k}},\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{1,k}},\frac{\partial\hat{\tau}}{\partial y}\right)^T, \quad (7)
\end{aligned}$$

where

$$\frac{\partial\hat{\tau}}{\partial\hat{\beta}_{0,i}}=\frac{d_i}{\hat{D}_{y,i}}\frac{1}{\hat{\beta}_{1,i}},\quad \frac{\partial\hat{\tau}}{\partial\hat{\beta}_{1,i}}=\frac{d_i}{\hat{D}_{y,i}}\frac{\log\frac{y}{1-y}-\hat{\beta}_{0,i}}{\hat{\beta}_{1,i}^2}$$

for $i=1,\ldots,k$, and

$$\frac{\partial\hat{\tau}}{\partial y}=-\frac{1}{y(1-y)}\left(\frac{1}{\hat{\beta}_{1,1}}\frac{d_1}{\hat{D}_{y,1}}+\cdots+\frac{1}{\hat{\beta}_{1,k}}\frac{d_k}{\hat{D}_{y,k}}\right),$$

Σ is the variance–covariance matrix of the $2k$ parameters $(\hat{\beta}_{0,1},\hat{\beta}_{1,1},\ldots,\hat{\beta}_{0,k},\hat{\beta}_{1,k})$ and the observed mean effect y at $(d_1,\ldots,d_k)$. Any two pairs of parameters, $(\hat{\beta}_{0,i},\hat{\beta}_{1,i})$ and $(\hat{\beta}_{0,j},\hat{\beta}_{1,j})$ when $i\neq j$ are independent since typically, different experimental subjects was used for drug i alone and for drug j alone, respectively. Further, all those subjects are different from the subjects administrated the combination dose $(d_1,\ldots,d_k)$. Thus, the estimates $(\hat{\beta}_{0,i},\hat{\beta}_{1,i})$ are independent of the estimates $(\hat{\beta}_{0,j},\hat{\beta}_{1,j})$ when $i\neq j$, and all of them are independent of the observed mean effect y at $(d_1,\ldots,d_k)$. Therefore, Σ is a block diagonal matrix with the block being a 2×2 matrix except for the last diagonal element $\operatorname{var}(y)$. An approximate variance of $\log(\hat{\tau})$ can be obtained by $\operatorname{var}(\log(\hat{\tau}))\simeq\frac{1}{\hat{\tau}^2}\operatorname{var}(\hat{\tau})$, where

$$\begin{aligned}
\operatorname{var}(\hat{\tau})\simeq{}&\sum_{i=1}^{k}\left(\frac{d_i}{\hat{D}_{y,i}}\right)^2\\
&\times\left(\frac{\operatorname{var}(\hat{\beta}_{0,i})}{\hat{\beta}_{1,i}^2}+\frac{2\operatorname{cov}(\hat{\beta}_{0,i},\hat{\beta}_{1,i})(\log\frac{y}{1-y}-\hat{\beta}_{0,i})}{\hat{\beta}_{1,i}^3}\right.\\
&\left.+\frac{\operatorname{var}(\hat{\beta}_{1,i})(\log\frac{y}{1-y}-\hat{\beta}_{0,i})^2}{\hat{\beta}_{1,i}^4}\right)\\
&+\left(\frac{1}{\hat{\beta}_{1,1}}\frac{d_1}{\hat{D}_{y,1}}+\cdots+\frac{1}{\hat{\beta}_{1,k}}\frac{d_k}{\hat{D}_{y,k}}\right)^2\\
&\times\left(\frac{1}{y(1-y)}\right)^2\operatorname{var}(y). \quad (8)
\end{aligned}$$

We can estimate $\operatorname{var}(y)$ in two ways. When there are replicates at the combination dose $(d_1,\ldots,d_k)$, $\operatorname{var}(y)$ can simply be estimated by the sample variance at (d_1,d_2). Otherwise, we may borrow the information from estimating the marginal dose–effect curves. Note that $\operatorname{var}(\log\frac{y}{1-y})\simeq(\frac{1}{y(1-y)})^2\operatorname{var}(y)$, thus, we may substitute $(\frac{1}{y(1-y)})^2\operatorname{var}(y)$ by the average of the squared residuals obtained from fitting the median-effect Equation (3) for all drugs involved assuming a constant variance for both the single and combination drug effects. Once the variance is obtained, a $(1-\alpha)\times 100\%$ confidence interval for $\log(\tau)$ can be formed as

$$\left[\log(\hat{\tau})-t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\log(\hat{\tau}))},\right.\qquad\left.\log(\hat{\tau})+t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\log(\hat{\tau}))}\right],$$

where $t_{n-2k,\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ percentile of t-distribution with $n-2k$ degree of freedom, and $n=\sum_{i=1}^{k}n_i$ with n_i $(i=1,\ldots,k)$ being the number of observations when drug i is used alone. $2k$ is the total number of estimated parameters involved in estimating the interaction index (6). Thus, a $(1-\alpha)\times 100\%$ confidence interval for τ can be constructed as:

$$\left[\hat{\tau}\exp\left(-t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\log(\hat{\tau}))}\right),\right.\qquad\left.\hat{\tau}\exp\left(t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\log(\hat{\tau}))}\right)\right]. \quad (9)$$

When $\operatorname{var}(\log(\hat{\tau}))$ is small, we have

$$\begin{aligned}
\hat{\tau}\exp&\left(\pm t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\log(\hat{\tau}))}\right)\\
&\simeq\hat{\tau}\exp\left(\pm\frac{t_{n-2k,\frac{\alpha}{2}}}{\hat{\tau}}\sqrt{\operatorname{var}(\hat{\tau})}\right)\\
&\simeq\hat{\tau}\pm t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\hat{\tau})}.
\end{aligned}$$

Therefore, if the error in (3) is small, the confidence interval for τ based on (9) is essentially the same as the confidence interval constructed by directly applying the delta method to $\hat{\tau}$, which is

$$\left[\hat{\tau}-t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\hat{\tau})},\quad\hat{\tau}+t_{n-2k,\frac{\alpha}{2}}\sqrt{\operatorname{var}(\hat{\tau})}\right]. \quad (10)$$

In Section 4, we illustrate that, for a large error in (3), the confidence interval (9) behaves better than (10) in two aspects: (i) the lower limit is greater than zero all the time; and (ii) the confidence interval has a coverage rate that is closer to the nominal rate. Therefore, the confidence interval (9) is preferred. When $n-2k$ is large, say $n-2k\geq 20$, one may use $z_{\frac{\alpha}{2}}$ instead of $t_{n-2k,\frac{\alpha}{2}}$ in estimating the confidence intervals (9) and (10), where $z_{\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ percentile of the standard normal distribution.

3. Interaction Indices and Their Confidence Bound at a Fixed Ray

3.1 Two-Drug Combination

Note that the confidence interval (9) is based on a single observation and the marginal dose–effect curves, and the estimated interaction index and its confidence interval are greatly influenced by this single observation. Chou and Talalay (1984) used a ray design to assess drug interactions. The advantage of their method is that it uses all observations with the component doses at a fixed ray. We adopt the interaction index instead of the combination index when using their approach. The basic idea (Chou 1991) is to regress $\log \frac{E}{1-E}$ on $\log D$ for each of the two drugs used alone and regress $\log \frac{E}{1-E}$ on $\log(d_1 + d_2)$ for the combination doses (d_1, d_2) with $\frac{d_2}{d_1} = \frac{\omega_2}{\omega_1}$, say, $\log \frac{E}{1-E} = \beta_{0,c} + \beta_{1,c} \log D_c$. Then for each fixed effect y, one may estimate the interaction index by

$$\hat{\tau}_{CT} = \frac{\hat{D}_{y,c}\frac{\omega_1}{\omega_1+\omega_2}}{\hat{D}_{y,1}} + \frac{\hat{D}_{y,c}\frac{\omega_2}{\omega_1+\omega_2}}{\hat{D}_{y,2}}, \tag{11}$$

where $\hat{D}_{y,1} = \exp(-\frac{\hat{\beta}_{0,1}}{\hat{\beta}_{1,1}}) \left(\frac{y}{1-y}\right)^{\frac{1}{\hat{\beta}_{1,1}}}$, $\hat{D}_{y,2} = \exp\left(-\frac{\hat{\beta}_{0,2}}{\hat{\beta}_{1,2}}\right) \left(\frac{y}{1-y}\right)^{\frac{1}{\hat{\beta}_{1,2}}}$, and $\hat{D}_{y,c} = \exp\left(-\frac{\hat{\beta}_{0,c}}{\hat{\beta}_{1,c}}\right) \left(\frac{y}{1-y}\right)^{\frac{1}{\hat{\beta}_{1,c}}}$. Commercial software *CalcuSyn* and *CompuSyn* are available for estimating the interaction indices and their confidence intervals. The confidence intervals for interaction indices in (11) were constructed based on Monte Carlo techniques and the normal assumption on the parameters (Belen'kii and Schinazi 1994). Briefly, the parameters, $(\hat{\beta}_{0,i}, \hat{\beta}_{1,i})$ for $i = 1, 2, c$, and the interaction index (11) are estimated from the observed data, then certain number of random samples (say, 500) of the parameters, $(\hat{\hat{\beta}}_{0,i}, \hat{\hat{\beta}}_{1,i})_j$ for $j = 1, \ldots, 500$ and $i = 1, 2, c$, are generated based on their estimated values and covariances under the normal assumption on each pair of these parameters. Thus, 500 interaction indices, $\hat{\hat{\tau}}_{CT,j}$ $(j = 1, \ldots, 500)$, can be calculated and its standard deviation can be estimated as $\hat{\sigma}_\tau^2 = \frac{1}{500} \sum_{j=1}^{500} (\hat{\hat{\tau}}_{CT,j} - \hat{\tau}_{CT})^2$. Consequently, the confidence interval can be constructed as $[\hat{\tau}_{CT} - z_{\frac{\alpha}{2}} \hat{\sigma}_\tau,\ \hat{\tau}_{CT} + z_{\frac{\alpha}{2}} \hat{\sigma}_\tau]$. In the simulation and case studies in Section 4, we used $t_{n_1+n_2+n_c-6, \frac{\alpha}{2}}$ instead of $z_{\frac{\alpha}{2}}$ since the number of observations is small.

In the following subsection, we extend Chou and Talalay's method to $k(\geq 2)$ drugs, estimate drug interaction at a fixed ray, say $d_1 : d_2 : \cdots : d_k = \omega_1 : \omega_2 : \cdots : \omega_k$, and construct a $(1-\alpha) \times 100\%$ confidence interval for the constructed interaction index at each effect y. Thus, by varying y, a pointwise confidence bound for the curve of interaction indices versus effects with combination doses at the fixed ray can be constructed by using the delta method.

3.2 *k*-drug combination

Again, we assume that the fitted dose–effect curve is $\log \frac{E}{1-E} = \hat{\beta}_{0,i} + \hat{\beta}_{1,i} \log D_i + \epsilon$ for drug i with $i = 1, \ldots, k$. The fitted dose–effect curve for the mixture with their component doses at a fixed ray with $d_1 : d_2 : \cdots : d_k = \omega_1 : \omega_2 : \cdots : \omega_k$ is $\log \frac{E}{1-E} = \hat{\beta}_{0,c} + \hat{\beta}_{1,c} \log D_c + \epsilon$ with $D_c = d_1 + d_2 + \cdots + d_k$. Then, for each fixed effect y, one may estimate interaction index by

$$\hat{\tau}_{CT} = \frac{\hat{D}_{y,c}\frac{\omega_1}{\omega_1+\ldots+\omega_k}}{\hat{D}_{y,1}} + \cdots + \frac{\hat{D}_{y,c}\frac{\omega_k}{\omega_1+\ldots+\omega_k}}{\hat{D}_{y,k}}, \tag{12}$$

where $\hat{D}_{y,i} = \left(\frac{y}{1-y}\right)^{\frac{1}{\hat{\beta}_{1,i}}} \exp\left(-\frac{\hat{\beta}_{0,i}}{\hat{\beta}_{1,i}}\right)$ for $i = 1, \ldots, k, c$. Again, $(\hat{\beta}_{0,i}, \hat{\beta}_{1,i})$ and $(\hat{\beta}_{0,j}, \hat{\beta}_{1,j})$ are independent as long as $i \neq j$ for $i, j = 1, \ldots, k, c$. Based on the delta method (Bickel and Doksum 2001), we can obtain an approximate variance for $\hat{\tau}_{CT}$

$$\begin{aligned} \operatorname{var}(\hat{\tau}_{CT}) &= \sum_{i=1}^{k} \left(\frac{\partial \hat{\tau}_{CT}}{\partial \hat{D}_{y,i}}\right)^2 \operatorname{var}(\hat{D}_{y,i}) \\ &\quad + \left(\frac{\partial \hat{\tau}_{CT}}{\partial \hat{D}_{y,c}}\right)^2 \operatorname{var}(\hat{D}_{y,c}) \\ &= \sum_{i=1}^{k} \left(-\frac{\omega_i \hat{D}_{y,c}}{(\sum_{i=1}^{k} \omega_i) \hat{D}_{y,i}^2}\right)^2 \operatorname{var}(\hat{D}_{y,i}) \\ &\quad + \left(\frac{1}{\sum_{i=1}^{k} \omega_i} (\sum_{i=1}^{k} \frac{\omega_i}{\hat{D}_{y,i}})\right)^2 \operatorname{var}(\hat{D}_{y,c}) \end{aligned} \tag{13}$$

with

$$\begin{aligned} \operatorname{var}(\hat{D}_{y,i}) &= \left(\frac{\partial \hat{D}_{y,i}}{\partial \hat{\beta}_{0,i}}, \frac{\partial \hat{D}_{y,i}}{\partial \hat{\beta}_{1,i}}\right) \Sigma_{\hat{\beta}_{0,i}, \hat{\beta}_{1,i}} \begin{pmatrix} \frac{\partial \hat{D}_{y,i}}{\partial \hat{\beta}_{0,i}} \\ \frac{\partial \hat{D}_{y,i}}{\partial \hat{\beta}_{1,i}} \end{pmatrix} \\ &= \hat{D}_{y,i}^2 \left(-\frac{1}{\hat{\beta}_{1,i}}, \frac{\hat{\beta}_{0,i} - \log \frac{y}{1-y}}{\hat{\beta}_{1,i}^2}\right) \\ &\quad \times \Sigma_{\hat{\beta}_{0,i}, \hat{\beta}_{1,i}} \begin{pmatrix} -\frac{1}{\hat{\beta}_{1,i}} \\ \frac{\hat{\beta}_{0,i} - \log \frac{y}{1-y}}{\hat{\beta}_{1,i}^2} \end{pmatrix} \end{aligned}$$

for $i = 1, \ldots, k, c$, respectively. Thus, replacing $\operatorname{var}(\hat{D}_{y,i})$ in (13), we can obtain the estimated variance for $\hat{\tau}_{CT}$:

$$\begin{aligned}\operatorname{var}(\hat{\tau}_{CT}) = {} & \sum_{i=1}^{k} \left(\frac{\omega_i \hat{D}_{y,c}}{(\sum_{i=1}^{k} \omega_i)\hat{D}_{y,i}} \right)^2 \\ & \times \left(\frac{\operatorname{var}(\hat{\beta}_{0,i})}{\hat{\beta}_{1,i}^2} + \frac{2\operatorname{cov}(\hat{\beta}_{0,i}, \hat{\beta}_{1,i})(\log \frac{y}{1-y} - \hat{\beta}_{0,i})}{\hat{\beta}_{1,i}^3} \right. \\ & \left. + \frac{\operatorname{var}(\hat{\beta}_{1,i})(\log \frac{y}{1-y} - \hat{\beta}_{0,i})^2}{\hat{\beta}_{1,i}^4} \right) \\ & + \left(\frac{\hat{D}_{y,c}}{\sum_{i=1}^{k} \omega_i} \left(\sum_{i=1}^{k} \frac{\omega_i}{\hat{D}_{y,i}} \right) \right)^2 \\ & \times \left(\frac{\operatorname{var}(\hat{\beta}_{0,c})}{\hat{\beta}_{1,c}^2} + \frac{2\operatorname{cov}(\hat{\beta}_{0,c}, \hat{\beta}_{1,c})(\log \frac{y}{1-y} - \hat{\beta}_{0,c})}{\hat{\beta}_{1,c}^3} \right. \\ & \left. + \frac{\operatorname{var}(\hat{\beta}_{1,c})(\log \frac{y}{1-y} - \hat{\beta}_{0,c})^2}{\hat{\beta}_{1,c}^4} \right). \end{aligned} \tag{14}$$

Here, we prefer the confidence interval based on the delta method on $\log(\hat{\tau}_{CT})$ since $\log(\hat{\tau}_{CT})$ is more approximately normally distributed than $\hat{\tau}_{CT}$. A $(1-\alpha) \times 100\%$ confidence interval for $\hat{\tau}_{CT}$ can be constructed by

$$\begin{aligned}\Bigg[& \hat{\tau}_{CT} \exp\left(\frac{-t_{n+n_C-2k-2,\frac{\alpha}{2}}}{\hat{\tau}_{CT}} \sqrt{\operatorname{var}(\hat{\tau}_{CT})} \right), \\ & \hat{\tau}_{CT} \exp\left(\frac{t_{n+n_C-2k-2,\frac{\alpha}{2}}}{\hat{\tau}_{CT}} \sqrt{\operatorname{var}(\hat{\tau}_{CT})} \right) \Bigg]. \end{aligned} \tag{15}$$

Again $n = \sum_{i=1}^{k} n_i$ and n_i $(i = 1, \ldots, k)$ is the number of observations when drug i is used alone, n_c is the number of observations for the combination doses at a fixed ray. By varying y in different values, we can construct a pointwise $(1-\alpha)100\%$ confidence bound for the curve of interaction indices versus effects. Thus, we can assess drug interactions for combination doses at the fixed ray while considering the stochastic uncertainty in obtaining the observations.

Remark: Comparing the variances of estimated interaction indices in (8) and (14), we note that the first terms in both equations are approximately the same, while the second terms are markedly different. In Sections 2 and 3, the k dose–effect curves for all the drugs involved are estimated. The first terms in both equations describe the uncertainty contributed by estimating the k marginal dose–effect curves. In Section 2, we estimate the interaction index based on the observed mean effect at a single combination dose, and oftentimes, we assume the combination dose is measured without error. Under this setting, the second term in (8) describes the variability contributed by the mean of the observed effects at the combination (d_1, d_2). In Section 3, we have the observations for combination doses at a fixed ray, then we fit the dose–effect curve for this ray. We estimate the interaction index for each fixed effect, where the combination dose producing such an effect is estimated. Thus, the second term in (14) describes the uncertainty contributed by the variance of the estimated combination dose $\hat{D}_{y,c}$, which could be split into the estimated combination dose

$$(\hat{d}_1, \ldots, \hat{d}_k) = \left(\frac{\omega_1}{\sum_{i=1}^{k} \omega_i} \hat{D}_{y,c}, \ldots, \frac{\omega_k}{\sum_{i=1}^{k} \omega_i} \hat{D}_{y,c} \right).$$

4. Simulations and Case Studies

4.1 Simulations

To examine whether the confidence intervals proposed in Sections 2 and 3 have proper characteristics, we performed simulations in the following two scenarios.

Scenario 1: three drugs, at a single combination dose. In the first scenario, we simulated three drugs that followed the median-effect Equation (2) with the same slope $m = -1$ and different median effective doses: $Dm_1 = 1$, $Dm_2 = 2$, and $Dm_3 = 4$, respectively. We took the combination dose (d_1, d_2, d_3) with each component being one third of its associated median effective dose, that is, $(d_1, d_2, d_3) = (\frac{Dm_1}{3}, \frac{Dm_2}{3}, \frac{Dm_3}{3}) = (\frac{1}{3}, \frac{2}{3}, \frac{4}{3})$. If the combination dose is additive, the expected effect will be 0.5. Let us denote the interaction index at this combination dose as τ, the effect as E, then based on Equation (4), we have

$$\frac{d_1}{Dm_1 \left(\frac{E}{1-E}\right)^{\frac{1}{m}}} + \frac{d_2}{Dm_2 \left(\frac{E}{1-E}\right)^{\frac{1}{m}}} + \frac{d_3}{Dm_3 \left(\frac{E}{1-E}\right)^{\frac{1}{m}}} = \tau.$$

Thus, the effect at (d_1, d_2, d_3) can be explicitly expressed as

$$E = \frac{\left(\tau^{-1}\left(\frac{d_1}{Dm_1} + \frac{d_2}{Dm_2} + \frac{d_3}{Dm_3}\right)\right)^m}{1 + \left(\tau^{-1}\left(\frac{d_1}{Dm_1} + \frac{d_2}{Dm_2} + \frac{d_3}{Dm_3}\right)\right)^m}.$$

We vary τ among (0.2, 0.4, 0.6, 0.8, 1, 1.25, 1.67, 2.5, 5). The corresponding effect E will be (0.167, 0.286, 0.375, 0.444, 0.5, 0.556, 0.625, 0.714, 0.833), respectively. Note that the slope m is negative, so the dose–effect curve is decreasing. If the effect is less than 0.5, then the combination dose will be synergistic and the interaction index will be less than 1; whereas if the effect is greater than 0.5, the combination dose will be antagonistic and the interaction index will be greater than 1. It is obvious that the farther the interaction index moves away from 1, the stronger is the interaction effect.

Table 1. Simulation results for Scenario 1: a fixed combination dose $(d_1, d_2, d_3) = (\frac{1}{3}, \frac{2}{3}, \frac{4}{3})$ but with varying interaction indices. The three dose–effect curves follow the median-effect equation with $m = -1$ and $(Dm_1, Dm_2, Dm_3) = (1, 2, 4)$.

	τ	0.2	0.4	0.6	0.8	1	1.25	1.67	2.5	5
$\sigma = 0.1$	mean($\hat{\tau}$)	0.202	0.402	0.601	0.804	1.001	1.266	1.676	2.503	5.022
	Cov.rate	0.950	0.952	0.946	0.947	0.939	0.958	0.951	0.947	0.948
	Cov.rate.log	0.954	0.955	0.950	0.950	0.941	0.953	0.951	0.951	0.954
	Len.ci	0.093	0.180	0.266	0.353	0.439	0.555	0.737	1.109	2.276
	Len.ci.log	0.094	0.181	0.268	0.357	0.443	0.560	0.744	1.119	2.298
	Pct.syn.log	100	100	99.4	51.7	3.4	0	0	0	0
	Pct.add.log	0	0	0.6	48.2	94.1	46.4	0.2	0	0
	Pct.ant.log	0	0	0	0.1	2.5	53.6	99.8	100	100
$\sigma = 0.4$	mean($\hat{\tau}$)	0.220	0.435	0.644	0.879	1.080	1.419	1.850	2.753	5.664
	Cov.rate	0.930	0.933	0.918	0.923	0.916	0.945	0.941	0.920	0.931
	Cov.rate.log	0.956	0.959	0.952	0.952	0.945	0.956	0.953	0.952	0.956
	Len.ci	0.397	0.777	1.151	1.562	1.926	2.552	3.381	5.108	10.966
	Len.ci.log	0.463	0.901	1.336	1.812	2.237	2.971	3.972	6.018	13.122
	Pct.syn.log	94.6	52.3	20.8	6.5	3.1	0.5	0.1	0	0
	Pct.add.log	5.4	47.7	79.2	92.2	94.5	91.1	79.2	48.4	5.9
	Pct.ant.log	0	0	0	1.3	2.4	8.4	20.7	51.6	94.1

Under the above setting, we first generated six equally spaced doses, ranging from 0.1 to three-fold of the associated median effective dose for each drug. We then generated the effects based on the model $\log \frac{E}{1-E} = \beta_0 + \beta_1 \log d + \epsilon$ with $\epsilon \sim N(0, \sigma^2)$ for each drug, where $\beta_0 = -m \log D_m$ and $\beta_1 = m$. We generated the observed effect at the combination dose $(\frac{1}{3}, \frac{2}{3}, \frac{4}{3})$ with the same size of the stochastic variation. The total sample size for each simulated experiment was 19 (six observations for each single drug and one observation for the combination dose effect). We fitted each dose–effect curve based on the generated data. Then, for each τ, we estimated the interaction index based on (6), constructed its 95% confidence intervals based on (9) and (10), respectively, calculated the length of the confidence intervals, and counted whether the true τ lies in the respective confidence intervals based on (9) and (10), and whether the confidence interval based on (9) lies below 1, contains 1, or lies above 1. We repeated this procedure 1,000 times, and averaged all the above quantities. We summarized the results in Table 1 under different settings for σ: $\sigma = 0.1$ and $\sigma = 0.4$, respectively. From Table 1, we conclude that (a) the estimation for τ (mean($\hat{\tau}$) in Table 1) is close to the true value and the accuracy decreases as σ increases; (b) the resulting coverage rates (Cov.rate.log) based on confidence interval (9) are closer to the nominal coverage rate of 95% than those (Cov.rate) based on (10), particularly, when σ is larger; (c) the average lengths of the confidence intervals (Len.ci.log) based on (9) and the average lengths of the confidence intervals (Len.ci) based on (10) increase as σ increases with Len.ci.log slightly larger than Len.ci to provide the nominal coverage rate; and (d) the percentage of times the model correctly assesses drug interaction based on (9) as synergy (Pct.syn.log), additivity (Pct.add.log), or antagonism (Pct.ant.log) decreases as σ increases. For each underlying interaction index among (0.2, 0.6, 1, 1.67, 5), we obtained the Q-Q plot of the 1,000 estimated interaction indices (Figure 2, Columns B1 and B3) as well as the Q-Q plot of the 1,000 logarithms of the estimated interaction indices (Figure 2, Columns B2 and B4) under the settings $\sigma = 0.1$ and $\sigma = 0.4$, respectively. From Figure 2, it is clear that for small σ (e.g., $\sigma = 0.1$), both the estimated interaction index and the logarithm of the estimated interaction index are approximately normally distributed. But, when σ becomes large, say $\sigma = 0.4$, the estimated interaction indices deviate from a normal distribution, while the logarithms of the estimated interaction indices are still approximately normally distributed. Therefore, one would expect that the delta method on the logarithm of the interaction index would work better for constructing confidence intervals for interaction indices. This assertion has been verified by the results of the current simulation studies. Thus, we prefer using confidence interval (9) over (10) for the interaction index in Section 2, and using the confidence interval (15) in Section 3.

Scenario 2: Two drugs, with a ray design. The second scenario involves two drugs that have the same dose–effect curves as drug 1 and drug 2 in the first scenario. That is, the two dose–effect curves follow the median-effect Equation (2) with the same slope $m = -1$, and median effective doses: $Dm_1 = 1$ and $Dm_2 = 2$, respectively. We assume that the dose–effect curve for the

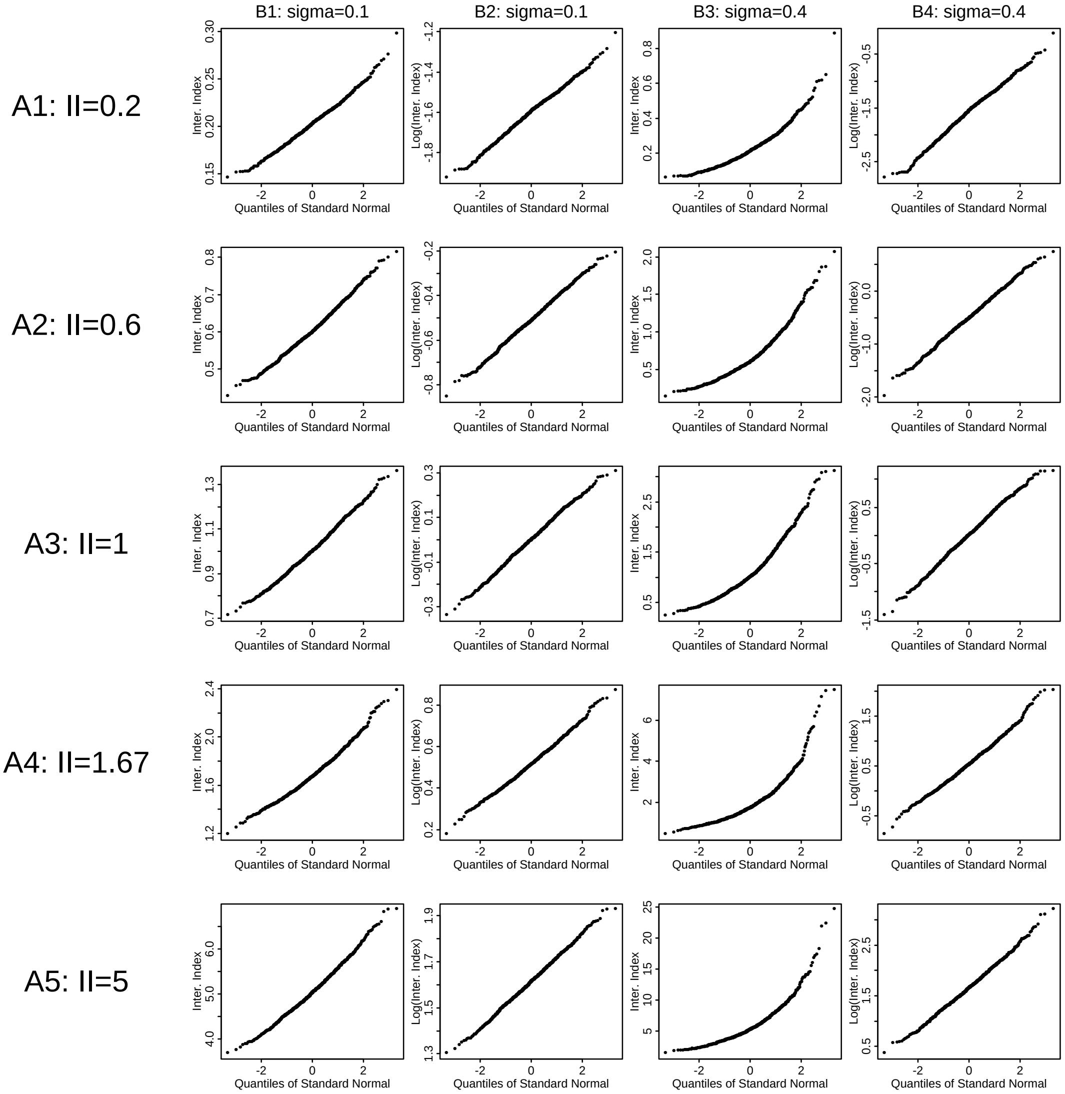


Figure 2. Q-Q plots of estimated interaction indices (Column B1 and B3) and Q-Q plots of the logarithms of the estimated interaction indices (Column B2 and B4) for 1000 samples under Scenario 1. Columns B1 and B2 show the Q-Q plots under $\sigma = 0.1$, and Column B3 and B4 show the Q-Q plots under $\sigma = 0.4$.

combination doses (d_1, d_2) at the fixed ray, say, $\frac{d_2}{d_1} = \frac{2}{1}$, follows the median-effect Equation (2) with $Dm_{12} = 1.5$ and $m_{12} = -2$.

We generated five equally spaced doses, ranging from 0.1 to three-fold of the associated median effective dose for each of the single drug, and five equally spaced doses, ranging from 0.5 to three-fold of the associated median effective dose for the mixture (d_1, d_2) at the fixed ray with $\frac{d_2}{d_1} = \frac{2}{1}$ and with the dose in the median effect Equation (2) being $d_1 + d_2$. We then generated the effects associated with these generated doses based on the model $\log \frac{E}{1-E} = \beta_0 + \beta_1 \log d + \epsilon$ with $\epsilon \sim N(0, \sigma^2)$ for the two drugs and their mixture under the settings $\sigma = 0.2$ and $\sigma = 0.4$, respectively. The total sample size for each simulated experiment is 15 (five observations for each single drug and five observations for combination doses). In addition, under each setting for σ, we generated seven samples for illustration. For each sample, we first fitted the dose–effect curves for the two drugs and the mixture, and then performed the following steps: we (a) estimated the interaction indices based on (12) for 42 equally spaced effect levels between the range of 0.1 to 0.95; (b) constructed their confidence intervals based on (15) and on a Monte Carlo simulation proposed by Belen'kii and Schinazi (1994), respectively; and (c) estimated interaction index (6) and constructed the confidence interval (9) for each observed combination dose. Figure 3 illustrates plots of the underlying curve of the interaction index versus effect (solid line), the pointwise 95% confidence bound for this curve based on (15) (dotted lines), the confidence bound based on the Monte Carlo simulation (dashed lines), and the estimated interaction indices (dots) and their confidence intervals (vertical bars) for the observed doses based on (9) for the seven illustrative samples under each setting for σ. From Figure 3, we conclude that (a) the pointwise 95% confidence bound (dotted lines) embrace the true curve (solid line) well; (b) the pointwise confidence bounds based on (15) (dotted lines) are similar to those based on Monte Carlo simulations (dashed lines) when σ is small, but perform better when σ is large; and (c) the confidence intervals based on single observations (vertical bars) are generally wider than both confidence bounds, and the conclusions based on confidence bounds are more accurate than those based on the confidence intervals (vertical bars) for single observations. In addition, for each sample, we calculated the ratio of the length of the confidence interval based on (9) and the confidence interval based on Monte Carlo simulation described in Section 3 for each of the observed effects at the combination doses at the fixed ray. Under $\sigma = 0.2$, the ratios for the seven samples have a mean 2.05 with standard deviation 0.58, and range from 1.12 to 3.28. Under $\sigma = 0.4$, the ratios for the seven samples have a mean 2.33 with standard deviation 0.92, and range from 1.01 to 4.15 after removing an extreme of the observation with effect close to one (Figure 3, Panel B7). Therefore, when several observations for combination doses at a fixed ray are available, the confidence bound derived in Section 3 using more available information is more efficient, thus preferred. In addition, when we ran this simulation to get the 14 confidence bounds in the 14 panels in Figure 3 by separately using the confidence interval based on (15) and Monte Carlo procedure on an Intel 1.83 GHz computer, the time it took was 10 seconds and 17 minutes, respectively. It is clear that the calculation for the confidence bound based on (15) is much faster than that based on Monte Carlo simulations.

4.2 Case Studies

Our research group (Lee et al. 2007; Kong and Lee 2006) investigated drug interactions between two novel agents, SCH66336 and 4-HPR, in a number of squamous cell carcinoma cell lines (Chun et al. 2003). Here we present the dataset and results from cell line UMSCC22B in Table 2 for investigating drug interaction in combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{1}$.

We first obtained the dose–effect curves for SCH66336 and 4-HPR by a linear regression of $\log \frac{E}{1-E}$ on $\log d$, based on the data in Table 2. Recall that $\log \frac{E}{1-E} = m(\log d - \log D_m) = \beta_0 + \beta_1 \log d$. The estimates of β_0, β_1, D_m, and $\hat{\sigma}$ for drug 1, drug 2, and the mixture of the drugs with equal concentrations are summarized in the same table.

The transformed data $\log \frac{E}{1-E}$ versus $\log d$ and the median-effect plots are shown in Figure 4(A). This median effect plot indicates that the data follow the median-effect Equation (2) reasonably well. Based on the fitted median-effect equations, we calculated the interaction indices based on (12) for varied effects for combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{1}$ and constructed their associated confidence bounds based on (15) and on Monte Carlo simulations (Belen'kii and Schinazi 1994), respectively. Figure 4(B) shows the plot of the interaction indices (on the logarithm scale) versus effects (solid line) for combination doses at this fixed ray and the 95% pointwise confidence bounds based on (15) (dotted lines) and on Monte Carlo simulations (dashed lines). Based on the confidence bound (dotted line), we conclude that the combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{1}$ with effect less than 0.52 are synergistic, and the combination doses at the fixed ray with effect greater than 0.52 are additive. The conclusions based on the confidence bounds obtained from Monte Carlo simulations (dashed lines) are slightly different. We also calculated four interaction indices based on (6) and their confidence intervals based on (9) at the four observed data points (d_1, d_2) as being (0.1, 0.1), (0.5, 0.5), (1, 1), and (2, 2). The four interaction indices were 0.791, 0.609, 0.256, and 0.103, and their corresponding 95% confidence intervals were [0.202, 3.091], [0.169, 2.193], [0.060, 1.087], and [0.018, 0.581], respectively. These pairs of interaction indices versus effects, along with their 95% confidence intervals are shown as vertical bars in Figure 4(B). From these four interaction indices and their confidence intervals, we conclude that the combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{1}$ are synergistic for doses ≥ 2 μM for each single drug, and additive for doses ≤ 1 μM for each single drug. The conclusions from the two procedures in Section 2 and 3 are slightly different: the combination doses (1, 1) and (0.5, 0.5) with respective observed effect 0.3551 and 0.4919 were identified as additive based on the second and third vertical bar (reading from left to right), while based on the confidence bound (dotted lines), the combination doses were identified as synergistic as each effect was less than 0.52. The relative length of the confidence interval based on (9) versus the monte Carlo confidence interval ranges from 1.27 to 3.72 for the four combination doses at the fixed ray. Once again, this example shows that the confidence interval estimation based on a single observation (9) is not as efficient as the corresponding confidence interval based on model (15) which used more data.

We also examined another dataset from a drug combi-

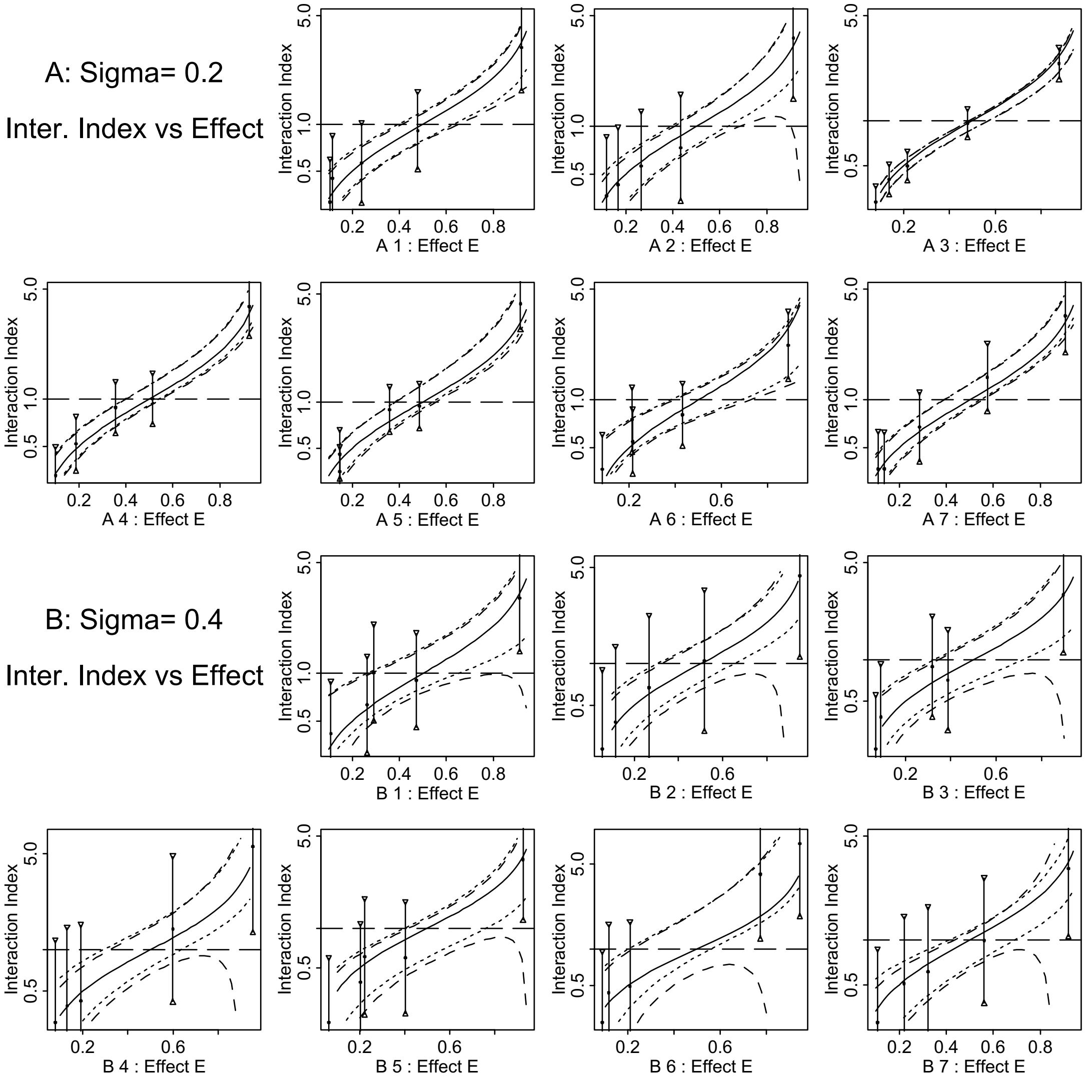


Figure 3. Simulated results under Scenario 2. The solid line is the plot of the underlying true interaction indices versus effects for combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{2}{1}$, the dotted lines and the dashed lines are the 95% pointwise confidence bounds for the curve of the interaction index versus effect based on the delta method and Monte Carlo simulation under the settings $\sigma = 0.2$ (Panels A) and $\sigma = 0.4$ (Panel B), respectively, and the dots and the vertical bars are the estimated interaction indices and their confidence intervals for observed combinations.

nation study for o-phenanthroline and ADP on the inhibition of horse liver alcohol dehydrogenase, which was analyzed by Chou and Talalay (1984) and by Belen'kii and Schinazi (1994). The dataset and the estimated median-effect parameters are shown in Table 3. The median-effect plots for the two drugs and their mixture (d_1, d_2) at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{17.4}$ are shown in Figure 5(A). The plot of interaction indices versus the fractional inhibitions at this fixed ray are shown as a solid line in Figure 5(B). In the same panel, we illustrate the pointwise confidence bound (dotted lines) based on (15) and the pointwise confidence bound (dashed line) based on Monte Carlo simulations for this curve, and the estimated interaction indices (dots) and their associated confidence intervals (vertical bars) based on (9) for the combination doses having observed effects. Again, the two 95% confidence bounds are almost the same, the vertical bars are wider than the confidence bound, and the conclusions on drug interactions based on vertical bars and those based on confidence bounds are consistent.

We developed two S-PLUS/R programs. One is used to estimate the interaction index and its confidence interval for a single combination dose of multiple drugs, and the other is used to estimate the pointwise

Table 2. Fractions of squamous cell carcinoma cells (UMSCC22B) surviving after 72 hours of treatment by single and combination dose levels of SCH66336 and 4-HPR and the fitted median-effect parameters.

SCH66336 dose (μM)	4-HPR dose dose (μM)	Fractional survival	Median-effect parameters
0.1		0.6701	$\hat{\beta}_{0,1} = 0.094(0.085)$
0.5		0.6289	$\hat{\beta}_{1,1} = -0.335(0.066)$
1		0.5577	$\hat{D}m_1 = 1.326$
2		0.4550	$\hat{\sigma}_1 = 0.187$
4		0.3755	
	0.1	0.7666	$\hat{\beta}_{0,2} = 0.217(0.073)$
	0.5	0.5833	$\hat{\beta}_{1,2} = -0.398(0.058)$
	1	0.5706	$\hat{D}m_2 = 1.726$
	2	0.4934	$\hat{\sigma}_2 = 0.129$
0.1	0.1	0.6539	$\hat{\beta}_{0,12} = -0.225(0.092)$
0.5	0.5	0.4919	$\hat{\beta}_{1,12} = -0.596(0.082)$
1	1	0.3551	$\hat{D}m_{12} = 0.686$
2	2	0.2341	$\hat{\sigma}_{12} = 0.182$

Note: The number inside the parentheses in the last column is the standard error of the estimate.

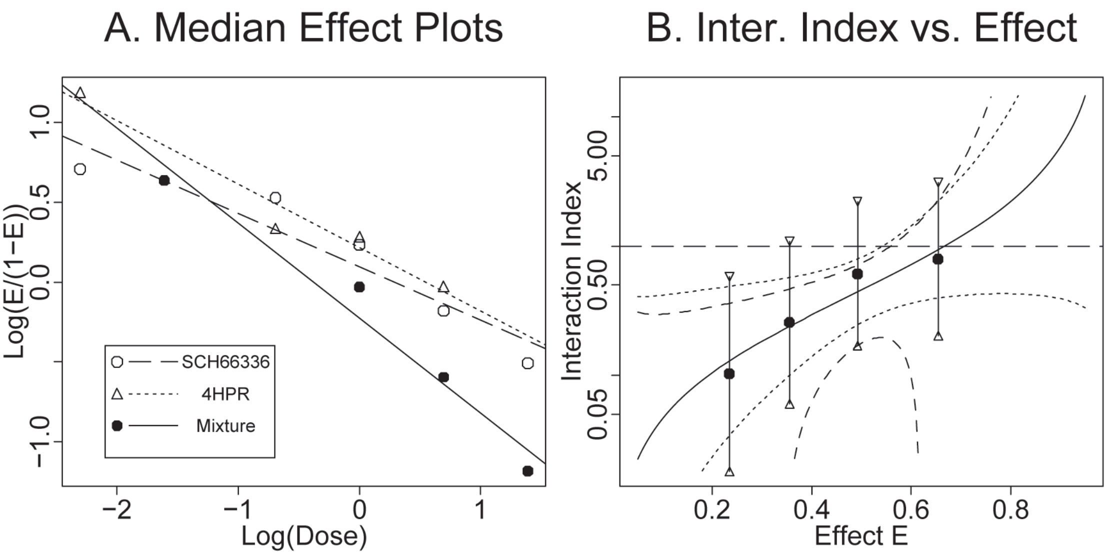


Figure 4. Median-effect plots (Panel A) and the plot of interaction indices versus effects (Panel B) for the combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{1}{1}$ for SCH66336 and 4HPR. In Panel B, the solid line is the plot of the estimated interaction indices versus effects, the two dotted lines and the two dashed lines are the pointwise 95% confidence bounds for the curve of interaction index versus effect based on the delta method in Section 3 and Monte Carlo simulation, respectively, and the dots and the vertical bars are the estimated interaction indices and their confidence intervals for observed combinations. The four vertical bars from left to right correspond to the combination doses of (2, 2), (1, 1), (0.5, 0.5), and (0.1, 0.1), respectively.

Table 3. Inhibition of horse liver alcohol dehydrogenase by o-phenanthroline and ADP alone and in combination (Chou and Talalay 1984; Belen’kii and Schinazi 1994).

o-phenanthroline	ADP	Fractional Inhibition	Median-effect parameters
8.7		0.132	$\hat{\beta}_{0,1} = -4.696(0.145)$
17.4		0.267	$\hat{\beta}_{1,1} = 1.302(0.046)$
26.1		0.411	$\hat{D}m_1 = 36.803$
34.8		0.476	$\hat{\sigma}_1 = 0.058$
43.5		0.548	
	0.5	0.175	$\hat{\beta}_{0,2} = -0.601(0.079)$
	1.0	0.400	$\hat{\beta}_{1,2} = 1.178(0.127)$
	1.5	0.492	$\hat{D}m_2 = 1.666$
	2.0	0.542	$\hat{\sigma}_2 = 0.161$
	2.5	0.592	
$9.2 \times \frac{17.4}{18.4}$	$9.2 \times \frac{1}{18.4}$	0.507	$\hat{\beta}_{0,12} = -3.843(0.038)$
$18.4 \times \frac{17.4}{18.4}$	$18.4 \times \frac{1}{18.4}$	0.769	$\hat{\beta}_{1,12} = 1.739(0.012)$
$27.6 \times \frac{17.4}{18.4}$	$27.6 \times \frac{1}{18.4}$	0.872	$\hat{D}m_{12} = 9.117$
$36.8 \times \frac{17.4}{18.4}$	$36.8 \times \frac{1}{18.4}$	0.919	$\hat{\sigma}_{12} = 0.015$
$46.0 \times \frac{17.4}{18.4}$	$46.0 \times \frac{1}{18.4}$	0.944	

Note: The number inside the parentheses in the last column is the standard error of the estimate.

confidence bound for the curve of interaction index versus effect for combination doses at a fixed ray. The S-PLUS/R code and the data example are available in *CI_of_Interaction_Index*, which can be downloaded from *http:// biostatistics.mdanderson.org/SoftwareDownload/*.

5. Discussion

We proposed a procedure in Section 2 to estimate the interaction index and constructed its associated confidence interval for a multiple drug combination. In most cases, the dose–effect for a single agent is known, and investigators are interested in assessing whether drug combinations are synergistic. When resources are limited, the experiment can be conducted in only a limited number of combination doses. We can assess drug interactions for those combination doses based on the procedure provided in Section 2. Note that although the dose–effect curves follow Chou and Talalay’s median-effect equation work reasonably well, the model may not work in certain cases. In these cases, other dose–effect models must be sought. For example, Lee et al. (2009) found that the $E_{\max}$ model describes the experimental data there better than Chou and Talalay’s median-effect equation; thus, the $E_{\max}$ model was used there. Upon finding the dose–effect curves of any parametric form which fits the data, one may use the same philosophy to estimate the interaction index and construct its associated confidence interval based on the delta method. However, using this “at a combination dose” method, one can assess drug interactions only at combination doses having observed effects, and the drug interaction tends to be predicted as additivity due to lack of efficiency (i.e., wide confidence intervals) even with nonadditive drug interactions.

Chou and Talalay’s method based on a ray design is widely used. We provided a procedure to construct pointwise confidence bound for Chou and Talalay’s curve of interaction index versus effect in Section 3. The procedure we provided avoids extensive calculations used in Monte Carlo techniques, which were required in the software *CalcuSyn* and *CompuSyn* and in the method provided by Belen’kii and Schinazi (1994). From the simulations and case studies in Section 4, we find that the confidence bounds provided in Section 3 are at least as good as the confidence bounds constructed using Monte Carlo techniques, while the confidence bounds in Section 3 are much faster to compute. Our limited simulation studies also show that the approximation based on the logarithm transformation and t-statistic works reasonably well when sample size was as low as 19 in one case and 15 in another case.

From simulations and case studies in Section 4, it is clear that the confidence intervals based on single observations (verticals bars in Figures 3, 4, and 5) are wider than the pointwise confidence bounds based on a ray design which use more data. In a ray design, the constructed confidence bound used all the information on this ray, therefore, the estimates based on Section 3 will be more

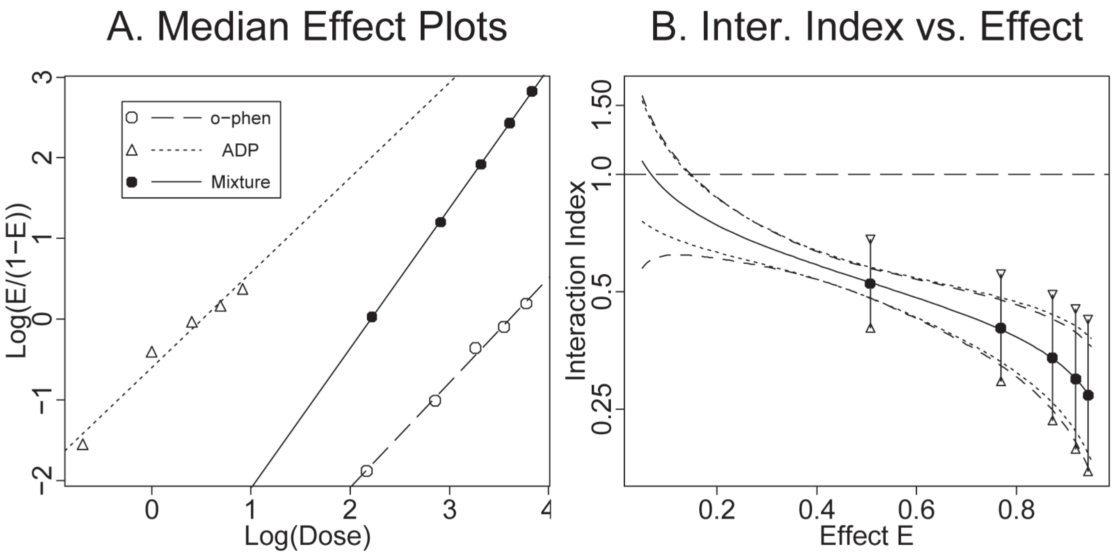


Figure 5. Median-effect plots (Panel A) and the plot of interaction indices versus effects (Panel B) for the combination doses at the fixed ray with $\frac{d_2}{d_1} = \frac{17.4}{1}$ for o-phenanthroline and ADP. In Panel B, the solid line is the plot of the estimated interaction indices versus effects, the two dotted lines and the two dashed lines are the pointwise confidence bounds for the curve of interaction index versus effect based on the delta method and Monte Carlo simulation, respectively, and the dots and the vertical bars are the estimated interaction indices and their confidence intervals for observed combinations.

efficient and more accurate. By examining the curve of interaction indices versus effects on several rays and examining their associated confidence bounds, one may obtain an overall picture of the drug interactions. The limitation is that one can only assess drug interactions for combination doses at these fixed examined rays. When a factorial design or a uniform design (Tan, Fang, Tian, and Houghton 2003) is used, a good strategy is to use response surface models, which use all the information presented in the observed data. We have proposed a generalized response surface model (Kong and Lee 2006) and a semiparametric model (Kong and Lee 2007) to capture drug interaction for all combination doses. However, response surface models for more than three drugs are difficult to construct. Therefore, to assess drug interactions among multiple drugs, the directly calculated interaction index and the plots of interaction indices versus effects at several fixed rays are still feasible and remain appealing methods to use. The confidence intervals we provided in Sections 2 and 3 are easy to calculate, and have a desirable coverage rate. Hence, it suggests that it is not necessary to run extensive Monte Carlo simulations for obtaining these confidence intervals. Based on the result of this article, Lee et al. (2009) constructed the simultaneous confidence interval for interaction indices over a range of treatment effects. The simultaneous confidence interval is also easy to calculate but is more conservative. The proposed confidence intervals can help us to gauge the uncertainties of the interaction indices for combination doses for two or more drugs and can also be used to provide more in-depth assessment for drug interactions.

Acknowledgments

The research was supported in part by grants from the National Cancer Institute CA106451, CA97007, CA16672, and Department of Defense W81XWH-04-1-0142 and W81XWH-05-2-0027. The authors are thankful to Lee Ann Chastain for editorial assistance, to the Editor, Associate Editor, and the referee for their constructive comments. The authors also thank Professor Paul Lewi for pointing out an error in an earlier version of the article.

[Received July 2006. Revised June 2007.]

References

Belen'kii, M. S., and Schinazi, R. F. (1994), "Multiple Drug Effect Analysis With Confidence Interval," *Antiviral Research*, 25, 1–11.

Berenbaum, M. C. (1985), "The Expected Effect of a Combination of Agents: The General Solution," *Journal of Theoretical Biology*, 114, 413–431.

——— (1989), "What is Synergy?," *Pharmacological Reviews*, 41, 93–141.

Bickel, P. J., and Doksum, K. A. (2001), *Mathematical Statistics: Basic Ideas and Selected Topics*, New Jersey: Prentice Hall, pp. 306–314.

Chou, T. C. (1991), "The Median-Effect Principle and the Combination Index for Quantitation of Synergism and Antagonism," in *Synergism and Antagonism in Chemotherapy*, eds. T. C. Chou and D. C. Rideout, San Diego: Academic Press, pp. 61–101.

——— (2006), "Theoretical Basis, Experimental Design, and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies," *Pharmacological Reviews*, 58, 621–681.

Chou, T. C., and Talalay, P. (1984), “Quantitative Analysis of Dose Effect Relationships: The Combined Effects of Multiple Drugs or Enzyme Inhibitors,” *Advances in Enzyme Regulation*, 22, 27–55.

Chun, K. H., Lee, J. J., Ayers, G. D., and Lotan, R. (2003), “Synergistic Induction of Apoptosis in Human Head and Neck Squamous Cell Carcinoma (HNSCC) Cell Lines by the Combination of the Synthetic Retinoid 4HPR and the Farnesyltransferase Inhibitor SCH66336,” *Proceedings of American Association for Cancer Research*, 44.

Greco, W. R., Bravo, G., and Parsons, J. C. (1995), “The Search of Synergy: A Critical Review from a Response Surface Perspective,” *Pharmacological Reviews*, 47(2), 331–385.

Kong, M., and Lee, J. J. (2006), “A General Response Surface Model with Varying Relative Potency for Assessing Drug Interactions,” *Biometrics*, 62(4), 986–995.

——— (2008), “A Semiparametric Response Surface Model for Assessing Drug Interactions,” *Biometrics*, 64, 396–405.

Lee, J. J., Kong, M., Ayers, G. D., and Lotan, R. (2007), “Interaction Index and Different Methods for Determining Drug Interaction in Combination Therapy,” *Journal of Biopharmaceutical Statistics*, 17, 461–480.

Lee, J. J., Lin, H. Y., Liu, D. D., and Kong, M. (2009), “Applying $E_{\max}$ Model and Interaction Index for Assessing Drug Interaction in Combination Studies,” *Frontiers in Biosciences*, in press.

Tan, M., Fang, H., Tian, G., and Houghton, P. J. (2003), “Experimental Design and Sample Size Determination for Testing Synergism in Drug Combination Studies Based on Uniform Measures,” *Statistics in Medicine*, 22, 2091–2100.

About the Authors

J. Jack Lee, Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, TX (E-mail: *jjlee@mdanderson.org*). Maiying Kong, Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, TX, and Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, KY (E-mail: *maiying.kong@louisville.edu*).

Oncogene (2008), 1–6
© 2008 Nature Publishing Group All rights reserved 0950-9232/08 $30.00
www.nature.com/onc

SHORT COMMUNICATION

Genomic profiling identifies *TITF1* as a lineage-specific oncogene amplified in lung cancer

KA Kwei[1], YH Kim[1], L Girard[2], J Kao[1], M Pacyna-Gengelbach[3], K Salari[4], J Lee[1], Y-L Choi[1,5], M Sato[2], P Wang[6], T Hernandez-Boussard[4], AF Gazdar[2], I Petersen[3], JD Minna[2] and JR Pollack[1]

[1]*Department of Pathology, Stanford University, Stanford, CA, USA;* [2]*Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, TX, USA;* [3]*Institute of Pathology, University Hospital Charité, Berlin, Germany;* [4]*Department of Genetics, Stanford University, Stanford, CA, USA;* [5]*Department of Pathology, Samsung Medical Center, Sungkyunkwan University, Seoul, South Korea and* [6]*Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA*

Lung cancer is a leading cause of cancer death, where the amplification of oncogenes contributes to tumorigenesis. Genomic profiling of 128 lung cancer cell lines and tumors revealed frequent focal DNA amplification at cytoband 14q13.3, a locus not amplified in other tumor types. The smallest region of recurrent amplification spanned the homeobox transcription factor *TITF1* (thyroid transcription factor 1; also called *NKX2-1*), previously linked to normal lung development and function. When amplified, *TITF1* exhibited increased expression at both the RNA and protein levels. Small interfering RNA (siRNA)-mediated knockdown of TITF1 in lung cancer cell lines with amplification led to reduced cell proliferation, manifested by both decreased cell-cycle progression and increased apoptosis. Our findings indicate that *TITF1* amplification and overexpression contribute to lung cancer cell proliferation rates and survival and implicate *TITF1* as a lineage-specific oncogene in lung cancer.
Oncogene advance online publication, 21 January 2008; doi:10.1038/sj.onc.1211012

Keywords: TITF1; lineage-specific oncogene; genomic profiling; lung cancer; TTF-1; NKX2-1

Lung cancer is the leading cause of cancer death in the United States (Jemal *et al.*, 2007). In lung cancers, the amplification of oncogenes such as *MYC*, *KRAS*, *MET*, *EGFR*, *ERBB2*, *CCND1* and *CDK4* contributes to tumor development and progression, and amplified genes have become important targets for molecularly-directed therapies (Sato *et al.*, 2007). To discover novel amplicons, we profiled 52 non-small cell lung cancer (NSCLC) cell lines and 76 NSCLC tumors (36 adenocarcinomas including 2 metastases, and 40 squamous cell carcinomas including 1 metastasis), by array-based comparative genomic hybridization (aCGH) on cDNA microarrays (Pollack *et al.*, 1999) covering ~22 000 human genes with a median inter-probe distance of ~30 kb. The most frequent focal DNA amplification not associated with a previously known oncogene occurred at cytoband 14q13.3 (Figure 1a).

Gain at 14q13.3 was present in 17 of 52 (33%) lung cancer cell lines, where it was more often observed in cell lines derived from adenocarcinomas (including bronchioloalveolar carcinomas) compared to other histologies ($P = 0.04$, Fisher's exact test; unspecified NSCLCs excluded from the analysis) (Figure 1b). 14q13.3 gain was also detectable in 4 of 36 (11%) lung adenocarcinomas and in 1 of 40 (3%) squamous cell carcinomas (all samples with gain were primary tumors). The lower frequencies observed in patient tumors may reflect an under-calling of gains due to contaminating non-cancerous cells in tumor samples or alternatively to a bias in the tumors attempted or successfully established as cultures or to selective pressures on cultured cells. Gain of 14q13.3 was significantly associated with the presence of *EGFR*-activating mutations ($P = 0.03$, Mann–Whitney *U*-test) (but not *KRAS* or *TP53* mutations), as well as the presence of specific DNA gains/losses elsewhere in the genome, including gain at 5p15.33 (*TERT*) (Supplementary Table 1). Notably, we have not observed the 14q13.3 locus to be amplified in other tumor types that we have profiled on the same platform, including cancers (totaling 385 specimens) of the breast, prostate, colon and pancreas (Bashyam *et al.*, 2005; Bergamaschi *et al.*, 2006; Lapointe *et al.*, 2007; unpublished data), suggesting that the putative driver oncogene(s) within this locus is lung cancer specific.

The smallest region of recurrent amplification, corroborated by CGH on a custom high-definition oligonucleotide microarray with probes spanning 14q13.2–q13.3 at 300 bp intervals (Figure 1c), included just eight named genes: *NFKBIA*, *INSM2*, *GARNL1*, *BRMS1L*, *MBIP*, *TITF1*, *NKX2-8* and *PAX9*. Because *TITF1* (thyroid transcription factor 1; also called *TTF-1* and *NKX2-1*) was known to participate in normal lung development (Kimura *et al.*, 1996) and had been characterized as a

---

Correspondence: Dr JR Pollack, Department of Pathology, Stanford University School of Medicine, CCSR-3245A, 269 Campus Drive, Stanford, CA 94305-5176, USA.
E-mail: pollack1@stanford.edu
Received 27 July 2007; revised 24 October 2007; accepted 1 December 2007

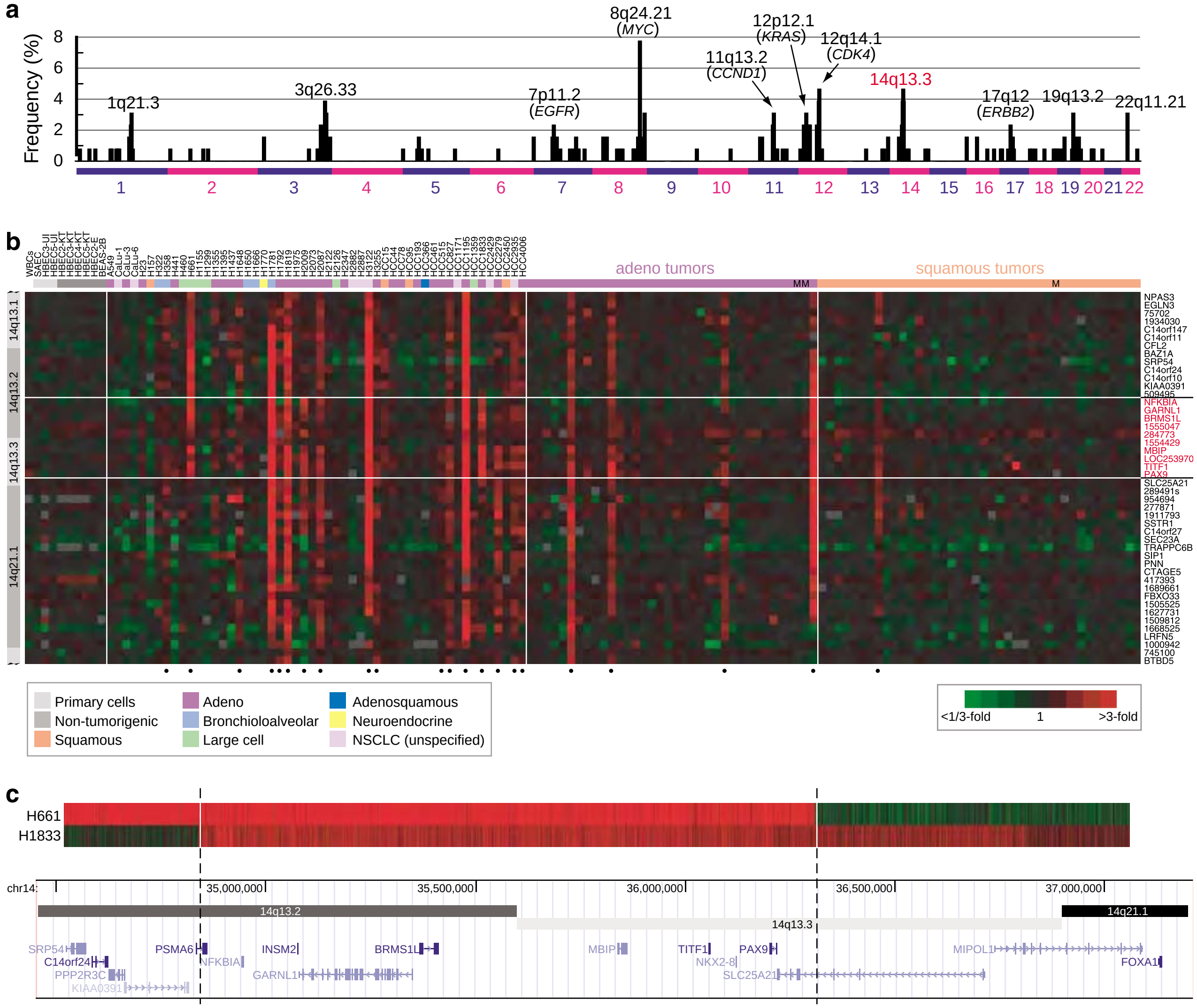


Figure 1 *TITF1* is focally amplified in lung cancer. (**a**) Frequency plot of cytobands harboring high-level DNA amplification in NSCLC cell lines and tumors (Supplementary Table 3). Cell lines were obtained from the Hamon Center for Therapeutic Oncology Research, UT Southwestern Medical Center. Tumors were banked at the University Hospital Charité, Berlin, Germany, with patient consent and Institutional Review Board approval, and DNA was extracted from several 30 μm cryostat tissue sections containing ⩾70% tumor cells. CGH was performed on cDNA microarrays from the Stanford Functional Genomics Facility containing 39 632 human cDNAs (representing 22 279 different mapped human genes/cDNAs), using our published protocol (Pollack *et al*., 2002). Map positions for arrayed cDNA clones were assigned using the NCBI genome assembly, accessed through the UCSC genome browser database (NCBI Build 36) (Kent *et al*., 2002). High-level DNA amplification was defined as tumor/normal aCGH ratios >3; selected cytobands with frequent amplification are indicated. The complete microarray data set is accessible from the GEO repository (GSE9995). (**b**) Genomic profiles by CGH on cDNA microarrays for NSCLC cell lines and tumors, histologies indicated (M = metastasis), for a segment of chromosome 14q13.1–q21.1. Genes are ordered by genome position. Red indicates positive tumor/normal aCGH ratios (scale shown), and samples called gained, using the fused lasso method (Tibshirani and Wang, 2008), at 14q13.3 are marked below by closed circles. Genes and ESTs (IMAGE clone ID shown) on the microarray residing within the amplicon core are highlighted by red text. (**c**) Genomic profiles by CGH on an Agilent (Santa Clara, CA, USA) high-definition custom microarray tiling 14q13.2–q21.1. The arrays comprised 10 614 probes tiling 3.3 Mb (nt 34 457 000–37 750 000) at 14q13.2–q21.1 with an average inter-probe spacing of 310 nt, with an additional 32 451 probes spanning the remaining genome for data normalization. DNAs were labeled as above, then hybridized to the array following the manufacturer's instructions. Arrays were scanned using an Agilent G2505B scanner and data extracted and normalized using Agilent Feature Extraction software (version 9.1) with default settings. Shown are two informative samples defining the amplicon boundaries, mapped onto the UCSC genome browser. The smallest region of recurrent amplification spans eight named genes. cDNA, complementary DNA; CGH, comparative genomic hybridization; NSCLC, non-small-cell lung cancer; UCSC, University of California Santa Cruz.

histological marker of lung adenocarcinoma (Travis *et al*., 2004), we sought to explore a possible functional connection of *TITF1* gene amplification with lung cancer.

Consistent with an oncogenic role, *TITF1* exhibited increased expression at both the RNA ($P = 0.046$, Mann–Whitney *U*-test) (Figure 2b) and protein level (Figure 2c) in NSCLC cell lines with amplification.

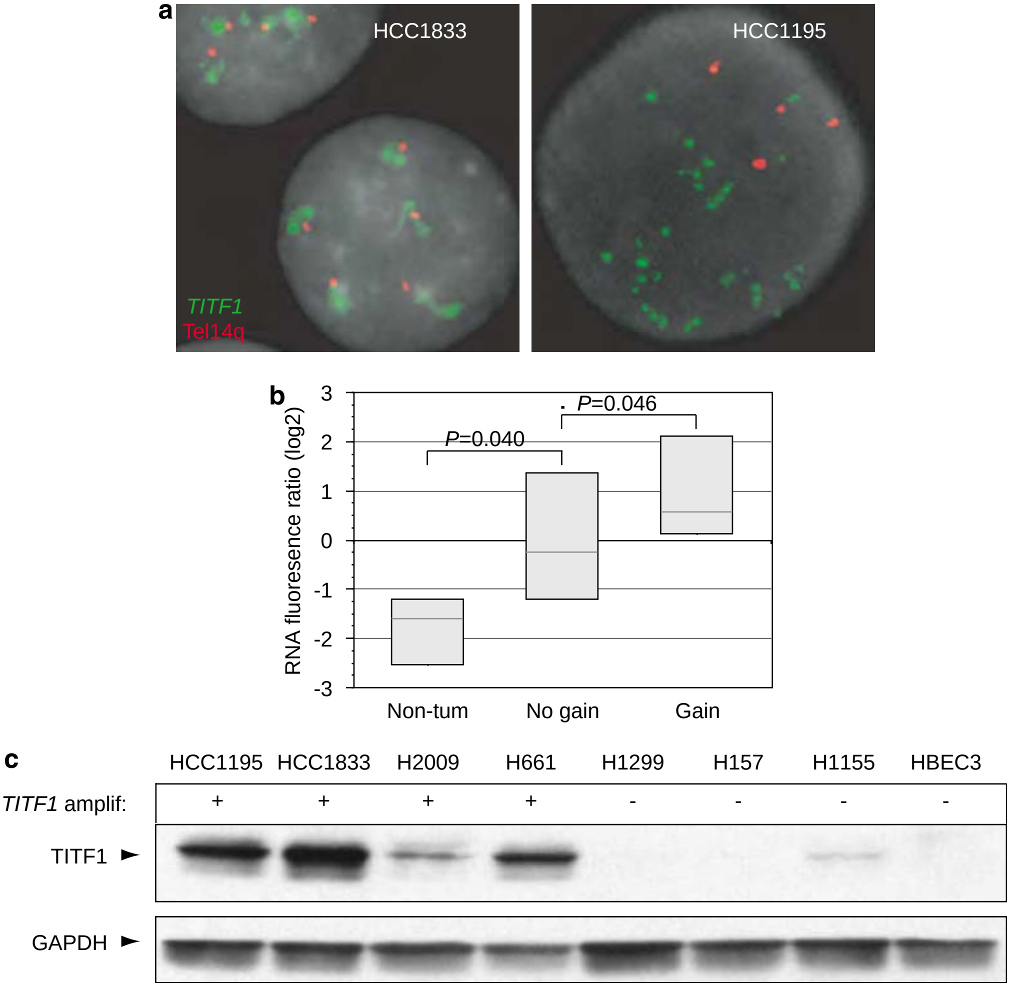
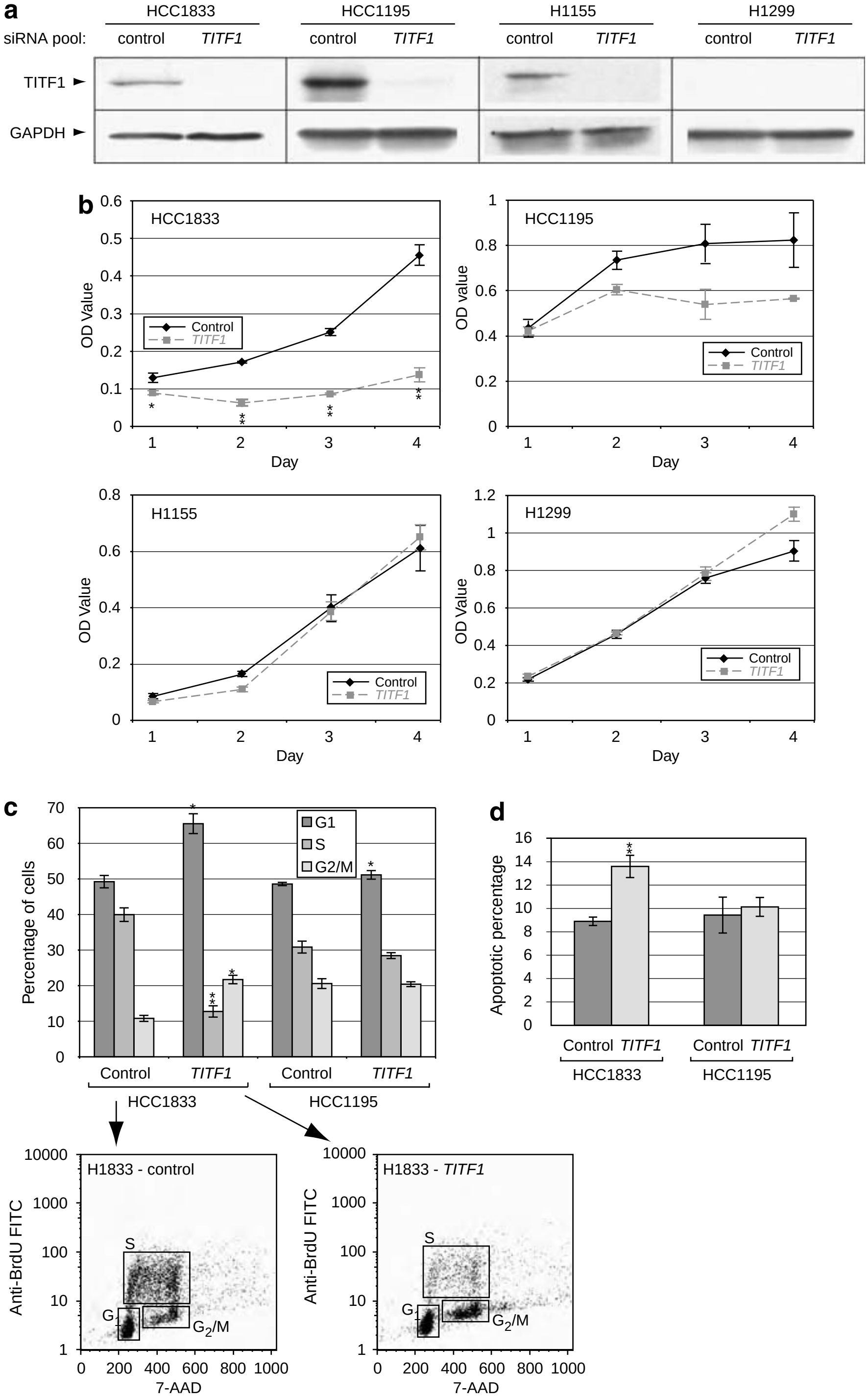


Figure 2 *TITF1* is overexpressed when amplified in lung cancer lines. (**a**) FISH validation of *TITF1* amplification in select NSCLC cell lines. FISH was performed using Vysis (Downers Grove, IL, USA) reagents according to the manufacturer's protocols. A locus-specific BAC mapping to *TITF1* at 14q13.3 (RP11-1083E2; BACPAC Resources Centre, Oakland, CA, USA) was labeled with SpectrumGreen, and co-hybridized with a SpectrumOrange-labeled telomere–14q probe (Vysis). Slides were counterstained with DAPI and imaged using an Olympus BX51 fluorescence microscope with Applied Imaging (San Jose, CA, USA) Cytovision 3.0 software. DNA amplification is evidenced by increased *TITF1* (green)/telomere–14q (red) signals (HCC1195, right) or by *TITF1* signal clusters (HCC1833, left). (**b**) mRNA levels of *TITF1*, measured by microarray, are elevated in NSCLC cell lines with *TITF1* amplification and also in comparison to primary and immortalized (but non-tumorigenic) lung epithelial cultures (Ramirez *et al*., 2004). Gene expression profiling was performed as described (Lapointe *et al*., 2004). Reported fluorescence ratios for TITF1 are normalized to the average TITF1 expression level across all samples. Box plots show 25th, 50th and 75th percentiles of expression. *P*-values (Mann–Whitney *U*-test) are indicated. (**c**) Western blot analysis of representative NSCLC cell lines indicates that *TITF1* is overexpressed at the protein level when amplified. Electrophoresis and blotting were performed as described (Kao and Pollack, 2006). TITF1 (~47 kDa) and GAPDH (loading control) were detected using anti-TITF1 rabbit polyclonal antibody (1:200; Santa Cruz Biotechnology, Santa Cruz, CA, USA) and anti-GAPDH rabbit polyclonal antibody (1:5000; Santa Cruz Biotechnology), followed by HRP-conjugated anti-rabbit IgG (1:20 000; Pierce, Rockford, IL, USA). Detection was carried out using the ECL kit (GE Healthcare, Piscataway, NJ, USA). FISH, fluorescence *in situ* hybridization; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; NSCLC, non-small-cell lung cancer; *TITF1*, thyroid transcription factor 1.

Notably, while other genes at 14q13.3 also exhibited increased expression when amplified, *TITF1* was the only well-measured gene that also exhibited significantly increased expression in tumorigenic compared to non-tumorigenic cell lines (Figure 2b and Supplementary Table 2). Sequencing of the *TITF1* open reading frame (and splice sites) from four NSCLC cell lines (HCC1195, HCC1833, H2009 and H661) with amplification revealed no DNA mutations, indicating that amplification-driven overexpression of the wild-type gene product would be of relevance.

To assess the functional significance of *TITF1* amplification and overexpression, we used RNA interference to target TITF1 knockdown in two lung cancer cell lines, HCC1833 and HCC1195, with *TITF1* amplification validated by fluorescence *in situ* hybridization (Figure 2a). Transfection of a Dharmacon On-TARGETplus pool of four different short interfering RNAs (siRNAs), designed and chemically modified to minimize off-target effects (Birmingham *et al*., 2006; Jackson *et al*., 2006), led to decreased TITF1 protein (Figure 3a) and decreased cell proliferation compared to a negative control siRNA pool (Figure 3b). Transfection of individual siRNAs from the same pool showed comparable effects (data not shown). In contrast, siRNA transfection of a lung cancer cell line (H1155) without *TITF1* amplification but with detectable expression did not diminish cell proliferation, indicating the functional importance of amplification-driven overexpression. Transfection of a cell line (H1299) without amplification or detectable expression also did not alter cell proliferation, supporting the specificity of *TITF1* targeting (Figure 3b). Similar negative results were observed upon transfection of a non-lung cancer cell line (colorectal cancer line SW48; data not shown). In the lung cancer cell lines with 14q13.3 amplification, the TITF1-targeted reduction in cell proliferation was attributable to both decreased cell-cycle progression (as evidenced by decreased S-phase fraction with G_1 block; Figure 3c) and increased apoptosis (Figure 3d). The effects were more pronounced in HCC1833 compared to HCC1195.



TITF1 is a tissue-specific transcription factor required for branching morphogenesis during normal lung development, as well as for the differentiation of pulmonary epithelial cells, as marked by the expression of surfactant proteins (which are among its transcriptional targets) (Bohinski *et al*., 1994; Kimura *et al*., 1996; Minoo *et al*., 1999). In the developing and adult lung, TITF1 is expressed mainly in peripheral airway cells and small-sized bronchioles (Yatabe *et al*., 2002). TITF1 has also been found to be expressed in approximately 40–50% of NSCLCs, more frequently in adenocarcinomas compared to squamous cell carcinomas and with expression linked to more favorable prognosis in some but not all studies (Puglisi *et al*., 1999; Tan *et al*., 2003). Immunostaining of TITF1 is used as the major lineage-specific marker to distinguish primary lung adenocarcinoma from metastatic adenocarcinoma to the lung (Travis *et al*., 2004). Our findings indicate that *TITF1* amplification and resultant overexpression contribute to increased cell proliferation rates and survival in lung cancer cells, and now implicate *TITF1* as a lung cancer-specific oncogene.

Recently, Tanaka *et al*. (2007) reported that TITF1 knockdown led to decreased colony formation, which they attributed to increased apoptosis in a cell line (NCI-H358) with TITF1 expression but (in their hands) no amplification. In our study, we showed that *TITF1* amplification led to protein overexpression and sensitivity to siRNA-mediated knockdown, highlighting the role of *TITF1* amplification as a critical event in the pathogenesis of a subset of lung cancers. Tanaka *et al*. also reported increased gene dosage of *TITF1*, measured by Southern blot and TaqMan PCR, in 2% of primary lung adenocarcinomas. The higher proportion we observed (11%) may reflect a higher sensitivity of aCGH (where multiple probes per locus are considered), or differences between the patient cohorts. Importantly, aCGH analysis also permitted us to define the 14q13 amplicon structure and boundaries, where our studies placed *TITF1* squarely within the amplicon core, consistent with a ‘driver’ role.

Given its connection to pulmonary epithelium differentiation, an oncogenic role of *TITF1* might seem counterintuitive. However, other tissue-specific transcription factors have been found to be amplified in cancers, including *MITF* in melanoma (Garraway *et al*., 2005), *AR* in prostate cancer (Visakorpi *et al*., 1995) and *ESR1* in breast cancer (Holst *et al*., 2007). The deregulated expression of such transcription factors, with normal roles in lineage proliferation or survival, may be required for tumor survival and progression in some cellular and genetic contexts, reflecting a state of ‘lineage dependency’ (Garraway and Sellers, 2006). More generally, the deregulated expression of transcription factors with roles in normal development reflects the principle of ‘oncology recapitulates ontogeny’ (Lechner *et al*., 2001).

Our finding that 14q13.3 amplification occurs mainly in lung adenocarcinomas (and their derived cell lines), the same histology in which TITF1 expression (even when not amplified) is predominantly restricted, is consistent with *TITF1* being the primary driver oncogene within 14q13.3. Nonetheless, other genes within the 14q13 amplicon may contribute to tumorigenesis. Notably, also residing within the amplicon core are two other homeodomain-containing genes, *NKX2-8* and *PAX9*; the former of which has been recently implicated in the control of normal lung development (Tian *et al*., 2006). Our preliminary data (not shown) indicate that *PAX9* is overexpressed when amplified in some NSCLC cell lines, and exhibits positive immunostaining in a subset of lung tumors. In another context, TITF1 and the paired-box member PAX8 have been shown to cooperatively activate the expression of thyroid-specific target genes (Miccadei *et al*., 2002). It is tempting to speculate that co-amplification of *NKX2-8* and/or *PAX9*, together with *TITF1*, may contribute to lung cancer development or progression. Indeed, very recently Kendall *et al*. (2007), who discovered the same 14q13.3 amplicon, reported that both *NKX2-8* and *PAX9* can synergize with *TITF1* to promote proliferation of immortalized human lung epithelial cells.

Future investigations are required to more precisely define the transcriptional effectors and pathways through which TITF1 mediates its oncogenic function. Nonetheless, our genomic profiling and functional studies implicate *TITF1* as a lineage-specific oncogene in lung cancer, a discovery that may lead to new opportunities for therapeutic intervention.

Figure 3 *TITF1* amplification/overexpression contributes to cell proliferation. (**a**) Confirmation of siRNA-mediated knockdown of target protein TITF1 by western blot. On-TARGETplus siRNAs targeting *TITF1*, along with a negative control siRNA pool (ON-TARGETplus siCONTROL Non-targeting Pool), were obtained from Dharmacon (Lafayette, CO, USA). Sequences of siRNAs are listed in Supplementary Table 4. Cell lines were maintained at 37 °C in RPMI-1640 with 10% fetal bovine serum. For transfection, 125 000–200 000 cells were seeded per well in a six-well plate and transfected using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol, using a final concentration of 50 nM siRNA for 6 h. Cell lysates were harvested 72 h post-transfection; GAPDH served as a loading control. (**b**) TITF1 knockdown results in decreased cell proliferation in cells with (HCC1833, HCC1195) but not without (H1155, H1299) *TITF1* amplification. At 24, 48, 72 and 96 h post-transfection, cell proliferation was quantified by colorimetry based on the metabolic cleavage of the tetrazolium salt WST-1 in viable cells, according to the manufacturer's protocol (Roche, Indianapolis, IN, USA). Transfections were performed in replicate and mean (± 1 s.d.) OD reported. (**c**) TITF1 knockdown reduces cell-cycle progression, evidenced by decreased S-phase fraction with G_1 block. At 72 h post-transfection, cell-cycle distribution analysis was performed using the BrdU-FITC Flow kit (BD Biosciences, San Jose, CA, USA) as per the manufacturer's instructions. Cells were incubated with 10 μM BrdU at 37 °C for 4 h prior to processing for analysis. Anti-BrdU FITC and 7-aminoactinomycin D (for total DNA content) stainings were scored by FACSCalibur (BD Biosciences) and analysed using CellQuest software (BD Biosciences). Transfections were performed in triplicate and mean (± 1 s.d.) cell-cycle fractions reported. Representative FACS plots are also shown. (**d**) TITF1 knockdown leads to increased apoptosis. At 72 h post-transfection, apoptosis was assayed by annexin-V staining and quantified by flow cytometry using the Vybrant Apoptosis Assay kit (Invitrogen) as per the manufacturer's instructions. Transfections were performed in triplicate, and mean (± 1 s.d.) percent apoptosis reported. $*P<0.05$. $**P<0.01$ (Student's *t*-test; *TITF1* compared to control). BrdU, bromodeoxyuridine; siRNA, small interfering RNA; *TITF1*, thyroid transcription factor 1.

Note added in proof: The *TITF1* amplicon was also recently identified by Weir *et al.* (2007).

Acknowledgements

We thank the SFGF for microarray manufacture, the SMD for database support, Ilana Galperin (Stanford Cytogenetics Laboratory) for assistance with FISH analysis and Eon Rios for assistance with FACS analysis. We also thank the members of the Pollack lab for helpful discussions. This work was supported in part by grants from the NIH: R01 CA97139 (JRP), T32 CA09151 (KAK), SPORE P50CA70907 (JDM), the DOD (JDM), the Vital, Anderson and Longenbaugh Foundations (JDM) and the Deutsche Krebshilfe: 10-2210-Pe4 (IP).

References

Bashyam MD, Bair R, Kim YH, Wang P, Hernandez-Boussard T, Karikari CA *et al.* (2005). Array-based comparative genomic hybridization identifies localized DNA amplifications and homozygous deletions in pancreatic cancer. *Neoplasia* **7**: 556–562.

Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE *et al.* (2006). Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* **45**: 1033–1040.

Birmingham A, Anderson EM, Reynolds A, Ilsley-Tyree D, Leake D, Fedorov Y *et al.* (2006). 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods* **3**: 199–204.

Bohinski RJ, Di Lauro R, Whitsett JA. (1994). The lung-specific surfactant protein B gene promoter is a target for thyroid transcription factor 1 and hepatocyte nuclear factor 3, indicating common factors for organ-specific gene expression along the foregut axis. *Mol Cell Biol* **14**: 5671–5681.

Garraway LA, Sellers WR. (2006). Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* **6**: 593–602.

Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S *et al.* (2005). Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**: 117–122.

Holst F, Stahl PR, Ruiz C, Hellwinkel O, Jehan Z, Wendland M *et al.* (2007). Estrogen receptor alpha (ESR1) gene amplification is frequent in breast cancer. *Nat Genet* **39**: 655–660.

Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J *et al.* (2006). Position-specific chemical modification of siRNAs reduces ‘off-target’ transcript silencing. *RNA* **12**: 1197–1205.

Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. (2007). Cancer statistics, 2007. *CA Cancer J Clin* **57**: 43–66.

Kao J, Pollack JR. (2006). RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes. *Genes Chromosomes Cancer* **45**: 761–769.

Kendall J, Liu Q, Bakleh A, Krasnitz A, Nguyen KC, Lakshmi B *et al.* (2007). Oncogenic cooperation and coamplification of developmental transcription factor genes in lung cancer. *Proc Natl Acad Sci USA* **104**: 16663–16668.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al.* (2002). The human genome browser at UCSC. *Genome Res* **12**: 996–1006.

Kimura S, Hara Y, Pineau T, Fernandez-Salguero P, Fox CH, Ward JM *et al.* (1996). The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev* **10**: 60–69.

Lapointe J, Li C, Giacomini CP, Salari K, Huang S, Wang P *et al.* (2007). Genomic profiling reveals alternative genetic pathways of prostate tumorigenesis. *Cancer Res* **67**: 8504–8510.

Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K *et al.* (2004). Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci USA* **101**: 811–816.

Lechner JF, Fugaro JM, Wong Y, Pass HI, Harris CC, Belinsky SA. (2001). Perspective: cell differentiation theory may advance early detection of and therapy for lung cancer. *Radiat Res* **155**: 235–238.

Miccadei S, De Leo R, Zammarchi E, Natali PG, Civitareale D. (2002). The synergistic activity of thyroid transcription factor 1 and Pax 8 relies on the promoter/enhancer interplay. *Mol Endocrinol* **16**: 837–846.

Minoo P, Su G, Drum H, Bringas P, Kimura S. (1999). Defects in tracheoesophageal and lung morphogenesis in Nkx2.1(−/−) mouse embryos. *Dev Biol* **209**: 60–71.

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF *et al.* (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23**: 41–46.

Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE *et al.* (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99**: 12963–12968.

Puglisi F, Barbone F, Damante G, Bruckbauer M, Di Lauro V, Beltrami CA *et al.* (1999). Prognostic value of thyroid transcription factor-1 in primary, resected, non-small cell lung carcinoma. *Mod Pathol* **12**: 318–324.

Ramirez RD, Sheridan S, Girard L, Sato M, Kim Y, Pollack J *et al.* (2004). Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res* **64**: 9027–9034.

Sato M, Shames DS, Gazdar AF, Minna JD. (2007). A translational view of the molecular pathogenesis of lung cancer. *J Thorac Oncol* **2**: 327–343.

Tan D, Li Q, Deeb G, Ramnath N, Slocum HK, Brooks J *et al.* (2003). Thyroid transcription factor-1 expression prevalence and its clinical implications in non-small cell lung cancer: a high-throughput tissue microarray and immunohistochemistry study. *Hum Pathol* **34**: 597–604.

Tanaka H, Yanagisawa K, Shinjo K, Taguchi A, Maeno K, Tomida S *et al.* (2007). Lineage-specific dependency of lung adenocarcinomas on the lung development regulator TTF-1. *Cancer Res* **67**: 6007–6011.

Tian J, Mahmood R, Hnasko R, Locker J. (2006). Loss of Nkx2.8 deregulates progenitor cells in the large airways and leads to dysplasia. *Cancer Res* **66**: 10399–10407.

Tibshirani R, Wang P. (2008). Spatial smoothing and hot spot detection for CGH data using the fused lasso. *Biostatistics* **9**: 18–29.

Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. (2004). *Pathology & Genetics: Tumours of the lung, pleura, thymus and heart. (World Health Organization Classification of Tumours)*. IARC Press: Lyon.

Visakorpi T, Hyytinen E, Koivisto P, Tanner M, Keinanen R, Palmberg C *et al.* (1995). *In vivo* amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* **9**: 401–406.

Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R *et al.* (2007). Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450**: 893–898.

Yatabe Y, Mitsudomi T, Takahashi T. (2002). TTF-1 expression in pulmonary adenocarcinomas. *Am J Surg Pathol* **26**: 767–773.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).

An overview of the molecular biology of lung cancer

Jill E. Larsen, Ph.D., Monica Spinola, Ph.D., Adi F. Gazdar, M.D., and John D. Minna, M.D.

Hamon Center for Therapeutic Oncology Research, Simmons Cancer Center

University of Texas Southwestern Medical Center

Correspondence:
John D. Minna, M.D.
Email: john.minna@utsouthwestern.edu
Telephone: 1-214-648-4900

This research was supported by:
National Cancer Institute Lung Cancer Specialized Program of Research Excellence (SPORE) (P50CA70907), Department of Defense VITAL and PROSPECT, NASA NSCOR (NNJ05HD36G), and by the Office of Science (BER) U.S. Department of Energy, Grant Number DE-AI02-05ER64068. JEL supported by TSANZ/Allen and Hanburys Respiratory Research Fellowship.

Introduction

Lung cancer cells have defects in the regulatory circuits that govern normal cell proliferation and homeostasis. Hanahan and Weinberg (1) described the “hallmarks of cancer” as six essential alterations in cell physiology that collectively dictate malignant growth. These acquired capabilities found in lung cancers are: self-sufficiency of growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis. Transformation from a normal to malignant lung cancer phenotype is thought to arise in a multi-step fashion, through a series of genetic and epigenetic alterations, ultimately evolving into invasive cancer by clonal expansion (2). These progressive pathological changes in the bronchial epithelium – known as preneoplastic or premalignant lesions – occur primarily as one of three distinct morphological forms: squamous dysplasia, atypical adenomatous hyperplasia, and diffuse idiopathic pulmonary neuroendocrine cell hyperplasia (3). Bronchial squamous dysplasia and carcinoma in situ (CIS) are the recognized preneoplastic lesions for squamous cell carcinoma; atypical adenomatous hyperplasia (AAH), a putative preneoplastic lesion, for a subset of adenocarcinomas; and neuroendocrine cell hyperplasia for neuroendocrine lung carcinomas (3). These preneoplastic lesions however account for the development of only a subset of lung cancers, for example, the precursor lesion for the most common neuroendocrine carcinoma of the lung, small cell lung carcinoma (SCLC), is unknown. Tumors are believed to become increasingly malignant with time, initiating tumorigenesis from possibly only a handful of mutations followed by additional (and different) mutations and epigenetic changes acquired during clonal expansion, where cells possessing *in vivo* growth advantage become dominant (2).

The identification and characterization of these molecular changes in lung cancer is of fundamental importance for improving the prevention, early detection, treatment and palliation of this disease. The overall goal is to translate these findings to the clinic by using molecular alterations as: 1) biomarkers for early detection; 2) targets for prevention; 3) tools for new molecular approaches; 4) signatures for personalizing prognosis and therapy selection for each patient, and 5) targets to specifically kill or inhibit the growth of lung cancer in patients.

Lung cancer arises from neoplastic changes to epithelial cells in the lung. However, it is not known whether all lung epithelial cells are susceptible to malignant transformation or only a subset of these cells, specifically; a major question is whether the changes need to take place in lung epithelial cells with stem cell-like properties. Lung cancer is a heterogeneous disease clinically, biologically, histologically and molecularly. The underlying causes of this heterogeneity are unknown and could reflect changes occurring in cells with various potential for differentiation (e.g. squamous or adenomatous) or represent different molecular changes occurring in the same target lung epithelial cells. This heterogeneity and molecular complexity contributes to the difficulty in unraveling the pathogenesis of lung cancer. Multiple oncogenes, tumor suppressor genes (TSGs), signaling pathway components and other cellular processes are involved in the molecular pathogenesis of lung cancer. This chapter will review molecular aberrations in lung cancer and the multiple pathways through which it develops.

The two main disease categories of lung cancer, non-small cell lung cancer (NSCLC) (representing 80-85% of cases) and small cell lung cancer (SCLC) (representing 15-20%) are generally classified based on differences in histological, clinical and neuroendocrine characteristics. NSCLC and SCLC can also differ molecularly with many genetic alterations exhibiting subtype specificity (summarized in Table 1). Additionally, molecular studies of

NSCLC have also revealed considerable differences between the subtypes of NSCLC, particularly the two most common subtypes: adenocarcinoma and squamous cell carcinoma.

Epidemiology and susceptibility in lung cancer

Eighty-five percent of lung cancers are caused by tobacco smoke where exposure to carcinogens present in tobacco smoke leads to the acquisition of genetic mutations that may eventually initiate carcinogenesis. However, not all lung cancers arise in smokers, and not all smokers will develop lung cancer. Thus, inherited factors must be involved which may predispose an individual to develop lung cancer – either by increasing susceptibility to the damaging effects of carcinogen exposure, or by increasing susceptibility regardless of smoking history. Worldwide, approximately 25% of lung cancer cases are not attributable to smoking (4). These cases occur more frequently in women, especially in Asiatic countries, target the distal airways, and are commonly adenocarcinomas. Coupled with molecular data that indicates strikingly different mutation patterns between known lung cancer genes such as *KRAS*, *EGFR*, and *TP53* and clinical data in relation to response to targeted therapies – it has now been suggested that lung cancer in never smokers be considered a distinct disease from the more common tobacco smoke-related lung cancer (4).

Many studies have examined the effect of single nucleotide polymorphisms (SNPs) on the risk of developing lung cancer (reviewed in (5, 6)). The reported risk effect in these studies is generally modest and often inconsistent, explaining why none are in routine use. However, meta-analyses as well as use of whole-genome SNP microarrays may hold the key to identifying robust and possible synergistic interactions between the modest affect of multiple SNPs. Of note, lung cancer risk was recently associated with genomic variation at 15q24/q25.1 by three separate studies simultaneously that used whole-genome SNP microarrays (7-9). While the conclusions of the three studies differed in whether the risk is conferred directly with cancer or through nicotine addiction, the genes within this locus – which include several genes encoding nicotinic acetylcholine receptor subunits (nAChR) – represent important targets for further functional analyses.

Genomic instability, telomeres and DNA damage in lung cancer

Malignant transformation is characterized by genetic instability which can exist at the chromosomal level (with loss or gain of genomic material, translocations, and microsatellite instability), at the nucleotide level (with single or several nucleotide base changes), or in the transcriptome (with altered gene expression). Abnormalities are typically targeted to proto-oncogenes, TSGs, DNA repair genes and other genes that can promote outgrowth of affected cells (10). The erosion of telomeres at the end of chromosomes is also associated with genomic instability leading to chromosomal abnormalities. Telomere length regulates the replicative capacity of a cell, where progressive telomere shortening occurs with each replication. Once the telomere becomes too short the cell will undergo cellular senescence or apoptosis. Activation of telomerase, the telomere-lengthening enzyme, in premalignant cells prevents loss of telomere ends beyond critical points and is essential for cell immortality. Although silenced in normal cells, telomerase is activated in >80% of NSCLCs and almost uniformly in SCLCs (11-13). In normal cells the presence of DNA damage engenders a DNA repair response, and if this is not successful, the apoptosis program is activated to remove the damaged cell. However, in premalignant and cancer cells the apoptosis program is often itself damaged thus allowing un- or mis-repaired DNA damage to persist in clones of cells.

Oncogenes and growth stimulatory pathways

Many oncogenes and TSGs have been identified by mapping of copy-number changes throughout the cancer genome (14-23). Earlier genomic analysis technology such as karyotyping and comparative genomic hybridization (CGH) enabled low resolution characterization of the lung cancer genome identifying whole arm/large-scale gain or loss on nearly every chromosomal arm, but most commonly 3p, 4q, 9p, and 17p loss and 1q, 3q, 5p, and 17q gain (24, 25) (Table 1). However, high resolution microarray analyses can now narrow in on these aberrant regions to detect focal amplifications and deletions often spanning only a handful of genes (Table 1).

Oncogenic activation typically occurs by gene amplification, point mutation, rearrangement, or through gene over-expression by other mechanisms including those mediated by miRNAs. These changes can result in persistent upregulation of mitogenic growth signals which induce cell growth. While promoting the malignant transformation of a cell, persistent upregulation of a particular growth signal or pathway can also result in 'oncogenic addiction' – whereby the cell becomes dependent upon this aberrant oncogenic signaling for survival (26). This presents an obvious target for therapeutics; remove or inhibit the oncogenic signal and an addicted tumor cell will die while normal 'non-addicted' cells will be unaffected. Signaling pathways commonly involved in lung cancer are shown in Figure 1.

Epidermal growth factor receptor signaling

The ErbB family of tyrosine kinase receptors includes four members – EGFR, ErbB-2 (HER2), ErbB-3, and ErbB-4 (27). While the intracellular tyrosine kinase domains of the four receptors are highly conserved the extracellular domain is not so conserved, enabling the receptors to bind different ligands. Following ligand binding, the ErbB receptors form homo- or heterodimers which results in receptor activation and subsequent activation of various signaling pathways. Activation of EGFR through the binding of EGF and EGF-like binding growth factors such as Transforming Growth Factor-α (TGF-α) enables the regulation of epithelial cell behavior and the initiation of tumors from epithelial cell origin through multiple signaling pathways. These include the RAS/RAF/MEK/MAPK pathway (cell proliferation), and the PI3K/AKT pathway and signal transduction and activator of transcription (STAT) 3 and STAT5 pathways (cell survival through the evasion of apoptosis) (28) (Figure 1). EGFR exhibits over-expression or aberrant activation in approximately 50-90% of NSCLCs with activating mutations occurring with or without amplification (29-32). Activating mutations, which are found with increased frequency in certain subsets of lung cancer patients, occur as three different types of somatic mutations – deletions, insertions, and missense point mutations – and are located in exons 19-21 which code for the tyrosine kinase domain of EGFR (17, 19). Mutant EGFRs (either by exon 19 deletion or exon 21 L858R mutation) show an increased amount and duration of EGFR activation compared with wildtype receptors (17), and have preferential activation of the PI3K/AKT and STAT3/STAT5 pathways rather than the RAS/RAF/MEK/MAPK pathway (33). EGFR mutant tumors (primarily adenocarcinomas) are initially highly sensitive to EGFR tyrosine kinase inhibitors (TKIs) (17-19). This represents an example of oncogene addiction in lung cancer where tumors initiated through EGFR mutation-activation of EGF signaling rely on continued EGF signaling for survival. However, despite an initial response, patients treated with EGFR TKIs eventually develop resistance to TKIs which is linked (in approximately 50% tumors) to the acquiring of a second mutation at T790M in exon 20 (34-39). Interestingly, the presence of the T790M mutation in a primary lung cancer that had not been treated with EGFR-

TKIs however, suggests that this resistance mutation may develop with tumor progression and not necessarily as a response to treatment (40). Recently, amplification of the MET proto-oncogene has been associated with acquired resistance to EGFR TKIs in 20% of resistant cases (36, 41) with MET activating the PI3K pathway through phosphorylation of ERBB3, independent of EGFR and ERBB2 (41). Importantly, inhibition of MET signaling was able to restore sensitivity to TKIs.

The RAS/RAF/MEK/MAPK/MYC pathway

The RAS proto-oncogene family (*KRAS*, *HRAS*, *NRAS* and *RRAS*) encode four highly homologous 21kDa membrane-bound proteins involved in signal transduction. Proteins encoded by the RAS genes exist in two states: an active state, in which GTP is bound to the molecule and an inactive state, where the GTP has been cleaved to GDP (42). Activating point mutations can confer oncogenic potential through a loss of intrinsic GTPase activity resulting in an inability to cleave GTP to GDP. This can initiate unchecked cell proliferation through the RAS/RAF/MEK/MAPK pathway, downstream of the EGFR signaling pathway (43). Activating *RAS* mutations occur in approximately 15-20% of NSCLCs and, in particular, 30-50% of adenocarcinomas (44). In lung cancer, 90% of mutations are located in *KRAS* (80% in codon 12, and the remainder in codons 13 and 61) with *HRAS* and *NRAS* mutations only occasionally documented (44). *KRAS* mutations are mutually exclusive with *EGFR* and *ERBB2* mutations, and confer resistance to EGFR TKIs and chemotherapy (45-47). Additionally, while *KRAS* mutations are primarily observed in lung adenocarcinomas of smokers, *EGFR* mutations are primarily observed in lung adenocarcinomas of never-smokers (4). These data demonstrate how lung adenocarcinoma can develop through different pathways, and it is likely, given the importance of EGFR targeted therapy that determination of *EGFR* and *KRAS* mutations in tumors will soon become part of standard care.

BRAF mutations occur in 1-5% of lung cancers and mutant BRAF mouse models can develop lung adenocarcinomas (48). The MYC proto-oncogene members are targets of RAS signaling and key regulators of numerous downstream pathways such as cell proliferation (49). Activation of MYC members often occurs through gene amplification. MYC is most frequently activated in NSCLC (50), with the other two members, MYCN and MYCL along with MYC, usually activated in SCLC (24, 51).

The PI3K/AKT pathway

The PI3K/AKT pathway which lies downstream of several receptor tyrosine kinases (RTKs) (such as EGFR) is a key regulator of cell proliferation, cell growth, and cell survival and is commonly activated in lung cancer through changes in several of its componenets including PI3K, PTEN, AKT, or EGFR or KRAS. In lung tumorigenesis, activation of the PI3K/AKT pathway is thought to occur early (52) and results in cell survival through inhibition of apoptosis. Activation can occur through the binding of the SH2-domains of p85, the regulatory subunit of PI3K, to phosphotyrosine residues of activated RTKs (53). Alternatively, activation can occur via binding of PI3K to activated RAS. Mutation and more commonly, amplification of *PIK3CA*, which encodes the catalytic subunit of phosphatidylinositol 3-kinase (PI3K), occurs most commonly in squamous cell carcinomas (20, 54-56). AKT, a serine/threonine kinase that acts downstream from PI3K can also have mutations that lead to pathway activation. One of the primary effectors of AKT is mTOR, a serine/threonine kinase involved in regulating proliferation, cell cycle progression, mRNA translation, cytoskeletal organization, and survival

(57). The tumor suppressor PTEN, which negatively regulates the PI3K/AKT pathway via phosphatase activity on phosphatidylinositol 3,4,5-trisphosphate (PIP3), a product of PI3K (58) is commonly suppressed in lung cancer by inactivating mutations or loss of expression (59, 60).

NKX2-1 (TITF1) – a lung cancer lineage dependent oncogene

Genome wide screens for DNA copy number changes in primary NSCLCs found multiple examples of amplification at 14q13.3 – and subsequent functional analysis (siRNA knockdowns in NSCLCs) identified *NKX2-1* (also termed *TITF1*) as the most likely target of amplification in lung cancer (14, 15, 61). *NKX2-1* encodes a lineage-specific transcription factor essential for branching morphogenesis in lung development and the formation of type II pneumocytes, the cells lining lung alveoli (62, 63). Amplification of tissue-specific transcription factors in cancer has been observed in *AR* in prostate cancer (64), *MITF* in melanoma (65), and *ESR1* in breast cancer (66). These findings have led to the development of a 'lineage-dependency' concept in tumors (67) whose survival and progression of a tumor is dependent upon continued signaling through a specific lineage pathways (i.e. abnormal expression of pathways involved in normal cell development) rather than continued signaling through the pathway of oncogenic transformation as seen with oncogene addiction (26).

EML4-ALK fusion proteins

Oncogenic fusion proteins created by recurrent chromosomal translocations are generally not common in solid tumors such as lung cancer; however, recent studies indicate this infrequency may be attributable to the difficulties in detection. The fusion of PTK echinoderm microtubule-associated protein like-4 (EML4)-anaplastic lymphoma kinase (ALK) was recently associated with lung cancer (68) and occurs in approximately 7% of NSCLCs (68-70). Fusing with EML4 induces a significant transforming potential in ALK. While wildtype ALK is thought to undergo transient homodimerization in response with specific ligand binding, EML4-ALK is constitutively oligomerized resulting in persistent mitogenic signaling and ultimately malignant transformation (71). Additionally, EML4-ALK generally appears to be mutually exclusive to that of EGFR or KRAS mutations in NSCLC and is more common in never or former smokers (72).

Tumor suppressor genes (TSGs) and growth inhibitory pathways

Loss of TSG function is an important step in lung carcinogenesis process and usually both alleles need to be inactivated. Generally, loss of heterozygosity (LOH) inactivates one allele through chromosomal deletion or translocation, and point mutation, epigenetic or transcriptional silencing inactivates the second allele (73, 74). In lung cancer, commonly inactivated TSGs include *TP53*, *RB1*, *CDKN2A*, *FHIT*, *RASSF1A* and *PTEN*.

The p53 pathway

TP53 (17p13) encodes a phosphoprotein which prevents accumulation of genetic damage in daughter cells. In response to cellular stress, p53 induces the expression of downstream genes such as cyclin-dependent kinase (CDK) inhibitors which regulate cell cycle checkpoint signals, causing the cell to undergo G1 arrest and allowing DNA repair or apoptosis (74). p53 inactivating mutations are the most common alterations in cancer, especially lung cancer, where 17p13 frequently demonstrates hemizygous deletion and mutational inactivation in the remaining allele (75-77). Regulation of p53 can occur through the oncogene MDM2, which reduces p53 levels through degradation, and the p14^{ARF} isoform of *CDKN2A*, which acts as a tumor suppressor by inhibiting MDM2. As such, the genes that encode MDM2 and p14^{ARF} are altered in lung cancer with amplification of *MDM2* seen in 6% of NSCLCs (78) and loss of p14^{ARF} expression in approximately 40% and 65% of NSCLCs and SCLCs, respectively (79, 80). Restoration of p53 expression *in vivo* has been achieved with p53 gene therapy of lung cancer patients in a subpopulation of tumor cells (81).

The CDKN2A/RB pathway

The CDKN2A-RB1 pathway controls G1 to S phase cell cycle progression. Hypophosphorylated retinoblastoma (RB) protein, encoded by *RB1*, halts the G1/S phase transition by binding to the transcription factor E2F1. This tumor suppressing effect can be inhibited by hyperphosphorylation of RB by CDK-CCND1 complexes (complexes between CDK4 or CDK6 and CCND1), and in turn, formation of CDK-CCND1 complexes can be inhibited by CDNK2A (82). Nearly all constituents of the CDKN2A/RB pathway have been shown to be altered in lung cancer through mutations (*CDK4* and *CDKN2A*), deletions (*RB1* and *CDKN2A*), amplifications (*CDK4* and *CCDN1*), methylation silencing (*CDKN2A* and *RB1*), and phosphorylation (RB) (83-88).

Chromosome 3p TSGs

Loss of one copy of chromosome 3p is one of the most frequent and early events in human cancer, found in 96% of lung tumors and 78% of lung preneoplastic lesions (89). Mapping of this loss identified several genes with functional tumor suppressing capacity including *FHIT* (3p14.2), *RASSF1A*, *TUSC2* (also called *FUS1*), and semaphorin family members *SEMA3B* and

SEMA3F (all at 3p21.3), and *RARβ* (3p24). In addition to LOH or allele loss, some of these 3p genes (*FHIT*, *RASSF1A*, *SEMA3B* and *RARβ*) often exhibit decreased expression in lung cancer cells by means of epigenetic mechanisms such as promoter hypermethylation (90-94). Additionally, *FHIT*, *RASSF1A*, *TUSC2*, and *SEMA3B* will reduce growth when re-introduced into lung cancer cells. *FHIT*, located in the most common fragile site in the human genome (*FRA3B*), has been shown to induce apoptosis in lung cancer (95). *RASSF1A* can induce apoptosis, as well as stabilize microtubules, and affect cell cycle regulation (96). The tumor suppressing effect of *TUSC2* is thought to occur via through inhibition of protein tyrosine kinases such as EGFR, PDGFR, c-Abl, c-Kit, and AKT (97) as well as inhibition of MDM2-mediated degradation of p53 (98). The candidate TSG *SEMA3B* encodes a secreted protein which can decrease cell proliferation and induce apoptosis when re-expressed in lung, breast and ovarian cancer cells (90, 91, 99, 100) in part, by inhibiting the AKT pathway (101). Another family member, *SEMA3F* may inhibit vascularization and tumorigenesis by acting on VEGF and ERK1/2 activation (102, 103) and *RARβ* exerts its tumor suppressing function by binding retinoic acid, thereby limiting cell growth and differentiation.

LKB1

The serine/threonine kinase LKB1 (also called STK11) is inactivated in ~30% of lung cancers and often correlates with KRAS activation (104), resulting in the promotion of cell growth. It functions as a TSG by regulating cell polarity, differentiation, and metastasis and can regulate cell metabolism (105). It has also been reported to inhibit the mTOR pathway (106).

Epigenetic regulation

Genetic abnormalities are associated with changes in DNA sequence however epigenetic events may lead to changes in gene expression without any changes in DNA sequence and therefore, importantly, the latter are potentially reversible (107). Aberrant promoter hypermethylation is an epigenetic change that occurs early in lung tumorigenesis and is found both in genes that normally undergo methylation in response to ageing, as well as in genes that normally remain unmethylated regardless of age (108). Gains of DNA methylation in a normally unmethylated promoter region of a gene results in silencing of gene transcription and is therefore a common method for the inactivation of tumor suppressor genes. In lung cancer, many genes have been found to be silenced by promoter hypermethylation (summarized in Table 1). They include genes involved in tumor suppression, tissue invasion, DNA repair, detoxification of tobacco carcinogens, and differentiation. Recent advances in whole-genome microarray profiling have allowed researchers to globally study DNA methylation patterns in lung cancer, the results of which have led to suggestions that the role of methylation in lung tumorigenesis has been underestimated (109-112). Restoration of expression of epigenetically silenced genes is a new targeted therapeutic approach. Histone deacetylation is an example of epigenetic change that can inhibit gene expression. Histone deacetylase inhibitors (HDACs) are being studied for the treatment of lung cancer and function by reversing gene silencing by inhibiting histone deacetylation (Figure 1 and Table 2).

MicroRNA-mediated regulation of lung cancer

MicroRNAs (miRNAs) are a recently identified class of non-protein encoding small RNAs present in the genomes of plants and animals. Ranging in size from 20-25 nucleotides, miRNAs are small RNA molecules that are capable of regulating gene expression by either direct cleavage of a targeted mRNA or inhibiting translation by interacting with the 3’ untranslated region (UTR) of a target mRNA. They are considered to play an important role in the pathogenesis of cancer – as either oncogenes or TSGs – due to abnormal expression found in several types of cancer, including lung cancer (113-121). Additionally, more than 50% of miRNAs are located in cancer-associated genomic regions or fragile sites (122, 123).

As observed for analyses on mRNA, protein and methylation patterns in lung cancer, miRNA microarrays have enabled the identification of many lung cancer-associated miRNAs (120, 121, 124-132). One of the most widely-studied miRNAs in lung cancer is the *let-7* miRNA family. Functioning as a tumor suppressor, it has been shown to regulate N-RAS, K-RAS and HMGA2 (133, 134) via binding to the let-7 binding sites in their respective 3’ UTRs (133, 135). It is frequently under-expressed in lung tumors, particularly NSCLC, compared to normal lung, and decreased expression has also been associated with poor prognosis (120, 125). Induction of *let-7* miRNA expression has been found to inhibit growth *in vitro* (120, 134, 136, 137) and reduce tumor development in a murine model of lung cancer (137, 138). In addition to *let*-7, other miRNAs with suggested tumor suppressing effects in lung cancer include *miR-126*, *miR-29a/b/c*, *miR-1* (125-128), and recently, *miR-128b* was reported to be a direct regulator of EGFR with frequent LOH occurring in NSCLC cell lines (129). Oncogenic miRNAs found to be over-expressed in lung cancer include the *miR-17-92* cluster of seven miRNAs (with suggested targets that include PTEN and RB), *miR-205*, *miR-21*, and *miR-155* (121, 130).

Lung cancer stem cells and Hedgehog, Notch and Wnt Signaling

The Hedgehog (HH), Wnt and Notch signaling pathways are important in normal lung development – specifically progenitor cell development and pulmonary organogenesis –

yhowever, they are now also being studied in regards to their role in tumor development (Figure 2). These signaling pathways are thought to be involved in the regulation on stem/progenitor cell self-renewal and maintenance, and while this process is normally tightly regulated process, genes that comprise these pathways are often mutated in human cancers (139-141), leading to abnormal activation of downstream effectors. In relation to cancer treatment, cancer stem cells are of great importance because they are thought to be resistant to cytotoxic therapies. If correct, this presents a need for effective therapies against these self-renewal signaling pathways.

In the HH pathway, increased signaling results in activation of the GLI oncogenes (GLI1, GLI2, and GLI3) that can regulate gene transcription (142-144). The HH signaling pathway was originally shown to have persistent activation in SCLC with high expression of SHH, PTCH, and GLI1 (145) but an important role in NSCLC was also recently demonstrated (146). The Notch signaling pathway is important in cell fate determination but can also promote and maintain survival in many human cancers (147-150). A recent study in mammary stem cells suggests the cytokine IL-6 may function as a regulator of self renewal in normal and tumor mammary stem cells through the Notch pathway through upregulation of the Notch-3 receptor (151), which is expressed in ~40% of resected lung cancers (152). The multifunctional cytokine IL-6 is involved in activation of JAK family of tyrosine kinases (153), which in turn activate multiple pathways through signaling molecules such as STAT3, MAPK, and PI3K (154). In lung adenocarcinomas, activated mutant EGFR has also been shown to induce levels of IL-6 leading to activation of STAT3 (155). The Wnt pathway has critical roles in organogenesis, cancer initiation and progression, and maintenance of stem cell pluripotency. In NSCLC, studies have found dysregulation of Wnt pathway members such as Wnt1, Wnt2 and Wnt7a, as well as upregulation of Wnt pathway agonists (Dvl proteins, LEF1, and Ruvb11) and underexpression or silencing of antagonists (WIF-1, sFRP1, CTNNBIP1, and WISP2) (156-162).

Human papilloma virus-mediated lung cancer

Human papilloma virus (HPV) has been identified in tumors from many organs, not just gynecological tumors. Nearly thirty years ago it was suggested to be a risk factor for lung cancer, particularly squamous cell carcinoma (163) and since then, many studies have investigated the role of HPV in lung cancer and have reported considerable geographical variation. A recent meta-analysis of 53 publications comprising 4,508 cases found the mean incidence of HPV in lung cancer was 25% and was detected in all subtypes of lung cancer, not just squamous cell (164). Studies from Europe and America had a lower incidence of 15-17% while Asian lung cancer cases reported a mean incidence of 38%. This observed high penetrance of HPV in lung cancer suggests more research is required to elucidate its role in lung cancer pathogenesis however, considering the significant variation observed between studies of cases from the same geographical location subsequent studies will need to have large sample and a detailed study design.

Conclusion

Genetic and epigenetic mechanisms underlying lung cancer development and progression continue to emerge. Over the past decade, research into the biology of many diseases has been spearheaded by the development of whole-genome microarray technology, allowing the simultaneous analysis of expression, copy-number and SNPs across thousands of genes. In lung cancer, gene expression studies have uncovered novel genes and pathways, as well as identified gene signatures that can better predict patient prognosis, response to treatment, and histology (165-168), reviewed (169, 170). High resolution mapping of alterations in the lung cancer genome has been able to identify single genes as targets of genomic gain or loss through improved definition of known aberrant regions or by identification of focal alterations undetectable with earlier technology (15, 171-173). Large-scale sequencing and SNP analyses have also led to the identification of novel somatic mutations or SNPs in the lung cancer genome (7-9, 174). While such genome-wide screens have the capacity of identifying novel genes or interactions in relation to lung cancer, the functional relevance of these findings still need to be elucidated using *in vitro* model systems such as tumor cell lines or immortalized human bronchial epithelial cells. These systems allow the characterization of single or sequential genetic alterations in relation to the development, maintenance, and progression of lung cancer and represent a crucial contribution in the understanding of the molecular biology of lung cancer. Functional characterization of genetic alterations and the signaling pathways with which they interact, has enabled the development of targeted therapies for the treatment of lung cancer (Table 2). Ranging from drugs in clinical use to those in clinical trial, they are directed against all known pathways of lung cancer initiation and progression such as proliferation, inhibition of apoptosis, angiogenesis, and invasion. This chapter has outlined some of the significant molecular alterations known to be involved in the initiation and/or progression of lung cancer. By characterizing these aberrations researchers endeavor to improve the detection, diagnosis, treatment, and prognosis of lung cancer through the integration of clinical and biological factors – to achieve personalized medicine.

References

1. Hanahan, D and Weinberg, RA. The hallmarks of cancer. *Cell* 2000;100: p. 57-70.
2. Nowell, PC. The clonal evolution of tumor cell populations. *Science* 1976;194: p. 23-8.
3. Wistuba, II and Gazdar, AF. Lung cancer preneoplasia. *Annu Rev Pathol* 2006;1: p. 331-48.
4. Sun, S, Schiller, JH, and Gazdar, AF. Lung cancer in never smokers--a different disease. *Nat Rev Cancer* 2007;7: p. 778-90.
5. Risch, A and Plass, C. Lung cancer epigenetics and genetics. *Int J Cancer* 2008;123: p. 1-7.
6. Herbst, RS, Heymach, JV, and Lippman, SM. Lung cancer. *N Engl J Med* 2008;359: p. 1367-80.
7. Amos, CI, Wu, X, Broderick, P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;40: p. 616-22.
8. Hung, RJ, McKay, JD, Gaborieau, V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008;452: p. 633-7.
9. Thorgeirsson, TE, Geller, F, Sulem, P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008;452: p. 638-42.
10. Sekido, Y, Fong, KM, and Minna, JD. Progress in understanding the molecular pathogenesis of human lung cancer. *Biochim Biophys Acta* 1998;1378: p. F21-59.
11. Albanell, J, Lonardo, F, Rusch, V, et al. High telomerase activity in primary lung cancers: association with increased cell proliferation rates and advanced pathologic stage. *J Natl Cancer Inst* 1997;89: p. 1609-15.
12. Hiyama, K, Hiyama, E, Ishioka, S, et al. Telomerase activity in small-cell and non-small-cell lung cancers. *J Natl Cancer Inst* 1995;87: p. 895-902.
13. Frias, C, Garcia-Aranda, C, De Juan, C, et al. Telomere shortening is associated with poor prognosis and telomerase activity correlates with DNA repair impairment in non-small cell lung cancer. *Lung Cancer* 2008;60: p. 416-25.
14. Kwei, KA, Kim, YH, Girard, L, et al. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. *Oncogene* 2008.
15. Weir, BA, Woo, MS, Getz, G, et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature* 2007;450: p. 893-8.
16. Merlino, GT, Xu, YH, Ishii, S, et al. Amplification and enhanced expression of the epidermal growth factor receptor gene in A431 human carcinoma cells. *Science* 1984;224: p. 417-9.
17. Lynch, TJ, Bell, DW, Sordella, R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004;350: p. 2129-39.
18. Pao, W, Miller, V, Zakowski, M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A* 2004;101: p. 13306-11.
19. Paez, JG, Janne, PA, Lee, JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004;304: p. 1497-500.
20. Massion, PP, Kuo, WL, Stokoe, D, et al. Genomic copy number analysis of non-small cell lung cancer using array comparative genomic hybridization: implications of the phosphatidylinositol 3-kinase pathway. *Cancer Res* 2002;62: p. 3636-40.

21. Friend, SH, Bernards, R, Rogelj, S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 1986;323: p. 643-6.
22. Nobori, T, Miura, K, Wu, DJ, Lois, A, Takabayashi, K, and Carson, DA. Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Nature* 1994;368: p. 753-6.
23. Steck, PA, Pershouse, MA, Jasser, SA, et al. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet* 1997;15: p. 356-62.
24. Fong, KM, Sekido, Y, and Minna, JD. Molecular pathogenesis of lung cancer. *J Thorac Cardiovasc Surg* 1999;118: p. 1136-52.
25. Fong, KM, Kida, Y, Zimmerman, PV, Ikenaga, M, and Smith, PJ. Loss of heterozygosity frequently affects chromosome 17q in non-small cell lung cancer. *Cancer Res* 1995;55: p. 4268-72.
26. Weinstein, IB. Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science* 2002;297: p. 63-4.
27. Normanno, N, Bianco, C, Strizzi, L, et al. The ErbB receptors and their ligands in cancer: an overview. *Curr Drug Targets* 2005;6: p. 243-57.
28. Arteaga, CL. Overview of epidermal growth factor receptor biology and its role as a therapeutic target in human neoplasia. *Semin Oncol* 2002;29: p. 3-9.
29. Rusch, V, Baselga, J, Cordon-Cardo, C, et al. Differential expression of the epidermal growth factor receptor and its ligands in primary non-small cell lung cancers and adjacent benign lung. *Cancer Res* 1993;53: p. 2379-85.
30. Franklin, WA, Veve, R, Hirsch, FR, Helfrich, BA, and Bunn, PA, Jr. Epidermal growth factor receptor family in lung cancer and premalignancy. *Semin Oncol* 2002;29: p. 3-14.
31. Herbst, RS. Review of epidermal growth factor receptor biology. *Int J Radiat Oncol Biol Phys* 2004;59: p. 21-6.
32. Fujino, S, Enokibori, T, Tezuka, N, et al. A comparison of epidermal growth factor receptor levels and other prognostic parameters in non-small cell lung cancer. *Eur J Cancer* 1996;32A: p. 2070-4.
33. Sordella, R, Bell, DW, Haber, DA, and Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 2004;305: p. 1163-7.
34. Pao, W, Miller, VA, Politi, KA, et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med* 2005;2: p. e73.
35. Kobayashi, S, Boggon, TJ, Dayaram, T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2005;352: p. 786-92.
36. Bean, J, Brennan, C, Shih, JY, et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci U S A* 2007;104: p. 20932-7.
37. Balak, MN, Gong, Y, Riely, GJ, et al. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin Cancer Res* 2006;12: p. 6494-501.
38. Kosaka, T, Yatabe, Y, Endoh, H, et al. Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib. *Clin Cancer Res* 2006;12: p. 5764-9.

39. Kwak, EL, Sordella, R, Bell, DW, et al. Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *Proc Natl Acad Sci U S A* 2005;102: p. 7665-70.
40. Kosaka, T, Yatabe, Y, Endoh, H, Kuwano, H, Takahashi, T, and Mitsudomi, T. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. *Cancer Res* 2004;64: p. 8919-23.
41. Engelman, JA, Zejnullahu, K, Mitsudomi, T, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007;316: p. 1039-43.
42. Mascaux, C, Iannino, N, Martin, B, et al. The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer* 2005;92: p. 131-9.
43. Shields, JM, Pruitt, K, McFall, A, Shaub, A, and Der, CJ. Understanding Ras: 'it ain't over 'til it's over'. *Trends Cell Biol* 2000;10: p. 147-54.
44. Rodenhuis, S and Slebos, RJ. Clinical significance of ras oncogene activation in human lung cancer. *Cancer Res* 1992;52: p. 2665s-2669s.
45. Pao, W, Wang, TY, Riely, GJ, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2005;2: p. e17.
46. Eberhard, DA, Johnson, BE, Amler, LC, et al. Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol* 2005;23: p. 5900-9.
47. Riely, GJ, Kris, MG, Rosenbaum, D, et al. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clin Cancer Res* 2008;14: p. 5731-4.
48. Ji, H, Wang, Z, Perera, SA, et al. Mutations in BRAF and KRAS converge on activation of the mitogen-activated protein kinase pathway in lung cancer mouse models. *Cancer Res* 2007;67: p. 4933-9.
49. Adhikary, S and Eilers, M. Transcriptional regulation and transformation by Myc proteins. *Nat Rev Mol Cell Biol* 2005;6: p. 635-45.
50. Nau, MM, Brooks, BJ, Jr., Carney, DN, et al. Human small-cell lung cancers show amplification and expression of the N-myc gene. *Proc Natl Acad Sci U S A* 1986;83: p. 1092-6.
51. Broers, JL, Viallet, J, Jensen, SM, et al. Expression of c-myc in progenitor cells of the bronchopulmonary epithelium and in a large number of non-small cell lung cancers. *Am J Respir Cell Mol Biol* 1993;9: p. 33-43.
52. West, KA, Linnoila, IR, Belinsky, SA, Harris, CC, and Dennis, PA. Tobacco carcinogen-induced cellular transformation increases activation of the phosphatidylinositol 3'-kinase/Akt pathway in vitro and in vivo. *Cancer Res* 2004;64: p. 446-51.
53. Vivanco, I and Sawyers, CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat Rev Cancer* 2002;2: p. 489-501.
54. Kawano, O, Sasaki, H, Endo, K, et al. PIK3CA mutation status in Japanese lung cancer patients. *Lung Cancer* 2006;54: p. 209-15.
55. Garnis, C, Lockwood, WW, Vucic, E, et al. High resolution analysis of non-small cell lung cancer cell lines by whole genome tiling path array CGH. *Int J Cancer* 2006;118: p. 1556-64.

56. Yamamoto, H, Shigematsu, H, Nomura, M, et al. PIK3CA mutations and copy number gains in human lung cancers. *Cancer Res* 2008;68: p. 6913-21.
57. Guertin, DA and Sabatini, DM. Defining the role of mTOR in cancer. *Cancer Cell* 2007;12: p. 9-22.
58. Maehama, T and Dixon, JE. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem* 1998;273: p. 13375-8.
59. Soria, JC, Lee, HY, Lee, JI, et al. Lack of PTEN expression in non-small cell lung cancer could be related to promoter methylation. *Clin Cancer Res* 2002;8: p. 1178-84.
60. Marsit, CJ, Zheng, S, Aldape, K, et al. PTEN expression in non-small-cell lung cancer: evaluating its relation to tumor characteristics, allelic loss, and epigenetic alteration. *Hum Pathol* 2005;36: p. 768-76.
61. Kendall, J, Liu, Q, Bakleh, A, et al. Oncogenic cooperation and coamplification of developmental transcription factor genes in lung cancer. *Proc Natl Acad Sci U S A* 2007;104: p. 16663-8.
62. Bingle, CD. Thyroid transcription factor-1. *Int J Biochem Cell Biol* 1997;29: p. 1471-3.
63. Ikeda, K, Clark, JC, Shaw-White, JR, Stahlman, MT, Boutell, CJ, and Whitsett, JA. Gene structure and expression of human thyroid transcription factor-1 in respiratory epithelial cells. *J Biol Chem* 1995;270: p. 8108-14.
64. Visakorpi, T, Hyytinen, E, Koivisto, P, et al. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* 1995;9: p. 401-6.
65. Garraway, LA, Widlund, HR, Rubin, MA, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* 2005;436: p. 117-22.
66. Holst, F, Stahl, PR, Ruiz, C, et al. Estrogen receptor alpha (ESR1) gene amplification is frequent in breast cancer. *Nat Genet* 2007;39: p. 655-60.
67. Garraway, LA and Sellers, WR. Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* 2006;6: p. 593-602.
68. Soda, M, Choi, YL, Enomoto, M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007;448: p. 561-6.
69. Meyerson, M. Cancer: broken genes in solid tumours. *Nature* 2007;448: p. 545-6.
70. Inamura, K, Takeuchi, K, Togashi, Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. *J Thorac Oncol* 2008;3: p. 13-7.
71. Mano, H. Non-solid oncogenes in solid tumors: EML4-ALK fusion genes in lung cancer. *Cancer Sci* 2008.
72. Koivunen, JP, Mermel, C, Zejnullahu, K, et al. EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. *Clin Cancer Res* 2008;14: p. 4275-83.
73. Knudson, AG, Jr. The ninth Gordon Hamilton-Fairley memorial lecture. Hereditary cancers: clues to mechanisms of carcinogenesis. *Br J Cancer* 1989;59: p. 661-6.
74. Breuer, RH, Postmus, PE, and Smit, EF. Molecular pathology of non-small-cell lung cancer. *Respiration* 2005;72: p. 313-30.
75. Takahashi, T, Nau, MM, Chiba, I, et al. p53: a frequent target for genetic abnormalities in lung cancer. *Science* 1989;246: p. 491-4.
76. Hollstein, M, Sidransky, D, Vogelstein, B, and Harris, CC. p53 mutations in human cancers. *Science* 1991;253: p. 49-53.

77. Greenblatt, MS, Bennett, WP, Hollstein, M, and Harris, CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* 1994;54: p. 4855-78.
78. Higashiyama, M, Doi, O, Kodama, K, et al. MDM2 gene amplification and expression in non-small-cell lung cancer: immunohistochemical expression of its protein is a favourable prognostic marker in patients without p53 protein accumulation. *Br J Cancer* 1997;75: p. 1302-8.
79. Gazzeri, S, Della Valle, V, Chaussade, L, Brambilla, C, Larsen, CJ, and Brambilla, E. The human p19ARF protein encoded by the beta transcript of the p16INK4a gene is frequently lost in small cell lung cancer. *Cancer Res* 1998;58: p. 3926-31.
80. Vonlanthen, S, Heighway, J, Tschan, MP, et al. Expression of p16INK4a/p16alpha and p19ARF/p16beta is frequently altered in non-small cell lung cancer and correlates with p53 overexpression. *Oncogene* 1998;17: p. 2779-85.
81. Ventura, A, Kirsch, DG, McLaughlin, ME, et al. Restoration of p53 function leads to tumour regression in vivo. *Nature* 2007;445: p. 661-5.
82. Ohtani, N, Yamakoshi, K, Takahashi, A, and Hara, E. The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression. *J Med Invest* 2004;51: p. 146-53.
83. Reissmann, PT, Koga, H, Takahashi, R, et al. Inactivation of the retinoblastoma susceptibility gene in non-small-cell lung cancer. The Lung Cancer Study Group. *Oncogene* 1993;8: p. 1913-9.
84. Merlo, A, Gabrielson, E, Askin, F, and Sidransky, D. Frequent loss of chromosome 9 in human primary non-small cell lung cancer. *Cancer Res* 1994;54: p. 640-2.
85. Brambilla, E, Moro, D, Gazzeri, S, and Brambilla, C. Alterations of expression of Rb, p16(INK4A) and cyclin D1 in non-small cell lung carcinoma and their clinical significance. *J Pathol* 1999;188: p. 351-60.
86. Sato, M, Takahashi, K, Nagayama, K, et al. Identification of chromosome arm 9p as the most frequent target of homozygous deletions in lung cancer. *Genes Chromosomes Cancer* 2005;44: p. 405-14.
87. Esteller, M. Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. *Adv Exp Med Biol* 2003;532: p. 39-49.
88. Kotake, Y, Cao, R, Viatour, P, Sage, J, Zhang, Y, and Xiong, Y. pRB family proteins are required for H3K27 trimethylation and Polycomb repression complexes binding to and silencing p16INK4alpha tumor suppressor gene. *Genes Dev* 2007;21: p. 49-54.
89. Wistuba, II, Behrens, C, Virmani, AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. *Cancer Res* 2000;60: p. 1949-60.
90. Ito, M, Ito, G, Kondo, M, et al. Frequent inactivation of RASSF1A, BLU, and SEMA3B on 3p21.3 by promoter hypermethylation and allele loss in non-small cell lung cancer. *Cancer Lett* 2005;225: p. 131-9.
91. Kuroki, T, Trapasso, F, Yendamuri, S, et al. Allelic loss on chromosome 3p21.3 and promoter hypermethylation of semaphorin 3B in non-small cell lung cancer. *Cancer Res* 2003;63: p. 3352-5.

92. Feng, Q, Hawes, SE, Stern, JE, et al. DNA methylation in tumor and matched normal tissues from non-small cell lung cancer patients. *Cancer Epidemiol Biomarkers Prev* 2008;17: p. 645-54.
93. Wistuba, II, Gazdar, AF, and Minna, JD. Molecular genetics of small cell lung carcinoma. *Semin Oncol* 2001;28: p. 3-13.
94. Zochbauer-Muller, S, Minna, JD, and Gazdar, AF. Aberrant DNA methylation in lung cancer: biological and clinical implications. *Oncologist* 2002;7: p. 451-7.
95. Siprashvili, Z, Sozzi, G, Barnes, LD, et al. Replacement of Fhit in cancer cells suppresses tumorigenicity. *Proc Natl Acad Sci U S A* 1997;94: p. 13771-6.
96. Agathanggelou, A, Cooper, WN, and Latif, F. Role of the Ras-association domain family 1 tumor suppressor gene in human cancers. *Cancer Res* 2005;65: p. 3497-508.
97. Ji, L and Roth, JA. Tumor suppressor FUS1 signaling pathway. *J Thorac Oncol* 2008;3: p. 327-30.
98. Deng, WG, Kawashima, H, Wu, G, et al. Synergistic tumor suppression by coexpression of FUS1 and p53 is associated with down-regulation of murine double minute-2 and activation of the apoptotic protease-activating factor 1-dependent apoptotic pathway in human non-small cell lung cancer cells. *Cancer Res* 2007;67: p. 709-17.
99. Tomizawa, Y, Sekido, Y, Kondo, M, et al. Inhibition of lung cancer cell growth and induction of apoptosis after reexpression of 3p21.3 candidate tumor suppressor gene SEMA3B. *Proc Natl Acad Sci U S A* 2001;98: p. 13954-9.
100. Ochi, K, Mori, T, Toyama, Y, Nakamura, Y, and Arakawa, H. Identification of semaphorin3B as a direct target of p53. *Neoplasia* 2002;4: p. 82-7.
101. Castro-Rivera, E, Ran, S, Brekken, RA, and Minna, JD. Semaphorin 3B inhibits the phosphatidylinositol 3-kinase/Akt pathway through neuropilin-1 in lung and breast cancer cells. *Cancer Res* 2008;68: p. 8295-303.
102. Brambilla, E, Constantin, B, Drabkin, H, and Roche, J. Semaphorin SEMA3F localization in malignant human lung and cell lines: A suggested role in cell adhesion and cell migration. *Am J Pathol* 2000;156: p. 939-50.
103. Kessler, O, Shraga-Heled, N, Lange, T, et al. Semaphorin-3F is an inhibitor of tumor angiogenesis. *Cancer Res* 2004;64: p. 1008-15.
104. Matsumoto, S, Iwakawa, R, Takahashi, K, et al. Prevalence and specificity of LKB1 genetic alterations in lung cancers. *Oncogene* 2007;26: p. 5911-8.
105. Alessi, DR, Sakamoto, K, and Bayascas, JR. LKB1-dependent signaling pathways. *Annu Rev Biochem* 2006;75: p. 137-63.
106. Shaw, RJ, Bardeesy, N, Manning, BD, et al. The LKB1 tumor suppressor negatively regulates mTOR signaling. *Cancer Cell* 2004;6: p. 91-9.
107. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;16: p. 6-21.
108. Baylin, SB, Esteller, M, Rountree, MR, Bachman, KE, Schuebel, K, and Herman, JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet* 2001;10: p. 687-92.
109. Rauch, TA, Zhong, X, Wu, X, et al. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proc Natl Acad Sci U S A* 2008;105: p. 252-7.
110. Shames, DS, Girard, L, Gao, B, et al. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. *PLoS Med* 2006;3: p. e486.

111. Zhong, S, Fields, CR, Su, N, Pan, YX, and Robertson, KD. Pharmacologic inhibition of epigenetic modifications, coupled with gene expression profiling, reveals novel targets of aberrant DNA methylation and histone deacetylation in lung cancer. *Oncogene* 2007;26: p. 2621-34.
112. Brena, RM, Morrison, C, Liyanarachchi, S, et al. Aberrant DNA methylation of OLIG1, a novel prognostic factor in non-small cell lung cancer. *PLoS Med* 2007;4: p. e108.
113. Metzler, M, Wilda, M, Busch, K, Viehmann, S, and Borkhardt, A. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. *Genes Chromosomes Cancer* 2004;39: p. 167-9.
114. Michael, MZ, SM, OC, van Holst Pellekaan, NG, Young, GP, and James, RJ. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res* 2003;1: p. 882-91.
115. Calin, GA, Dumitru, CD, Shimizu, M, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 2002;99: p. 15524-9.
116. Eis, PS, Tam, W, Sun, L, et al. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci U S A* 2005;102: p. 3627-32.
117. He, L, Thomson, JM, Hemann, MT, et al. A microRNA polycistron as a potential human oncogene. *Nature* 2005;435: p. 828-33.
118. Ota, A, Tagawa, H, Karnan, S, et al. Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. *Cancer Res* 2004;64: p. 3087-95.
119. Voorhoeve, PM, le Sage, C, Schrier, M, et al. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell* 2006;124: p. 1169-81.
120. Takamizawa, J, Konishi, H, Yanagisawa, K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* 2004;64: p. 3753-6.
121. Hayashita, Y, Osada, H, Tatematsu, Y, et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res* 2005;65: p. 9628-32.
122. Calin, GA, Sevignani, C, Dumitru, CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* 2004;101: p. 2999-3004.
123. Sevignani, C, Calin, GA, Nnadi, SC, et al. MicroRNA genes are frequently located near mouse cancer susceptibility loci. *Proc Natl Acad Sci U S A* 2007;104: p. 8017-22.
124. Yu, SL, Chen, HY, Chang, GC, et al. MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* 2008;13: p. 48-57.
125. Yanaihara, N, Caplen, N, Bowman, E, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 2006;9: p. 189-98.
126. Fabbri, M, Garzon, R, Cimmino, A, et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc Natl Acad Sci U S A* 2007;104: p. 15805-10.
127. Crawford, M, Brawner, E, Batte, K, et al. MicroRNA-126 inhibits invasion in non-small cell lung carcinoma cell lines. *Biochem Biophys Res Commun* 2008;373: p. 607-12.

128. Nasser, MW, Datta, J, Nuovo, G, et al. Downregulation of microRNA-1 (miR-1) in lung cancer: Suppression of tumorigenic property of lung cancer cells and their sensitization to doxorubicin induced apoptosis bymiR-1. *J Biol Chem* 2008.
129. Weiss, GJ, Bemis, LT, Nakajima, E, et al. EGFR regulation by microRNA in lung cancer: correlation with clinical response and survival to gefitinib and EGFR expression in cell lines. *Ann Oncol* 2008;19: p. 1053-9.
130. Markou, A, Tsaroucha, EG, Kaklamanis, L, Fotinou, M, Georgoulias, V, and Lianidou, ES. Prognostic Value of Mature MicroRNA-21 and MicroRNA-205 Overexpression in Non-Small Cell Lung Cancer by Quantitative Real-Time RT-PCR. *Clin Chem* 2008.
131. Garofalo, M, Quintavalle, C, Di Leva, G, et al. MicroRNA signatures of TRAIL resistance in human non-small cell lung cancer. *Oncogene* 2008;27: p. 3845-55.
132. Volinia, S, Calin, GA, Liu, CG, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* 2006;103: p. 2257-61.
133. Johnson, SM, Grosshans, H, Shingara, J, et al. RAS is regulated by the let-7 microRNA family. *Cell* 2005;120: p. 635-47.
134. Lee, YS and Dutta, A. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev* 2007;21: p. 1025-30.
135. Mayr, C, Hemann, MT, and Bartel, DP. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* 2007;315: p. 1576-9.
136. Johnson, CD, Esquela-Kerscher, A, Stefani, G, et al. The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res* 2007;67: p. 7713-22.
137. Esquela-Kerscher, A, Trang, P, Wiggins, JF, et al. The let-7 microRNA reduces tumor growth in mouse models of lung cancer. *Cell Cycle* 2008;7: p. 759-64.
138. Kumar, MS, Erkeland, SJ, Pester, RE, et al. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proc Natl Acad Sci U S A* 2008;105: p. 3903-8.
139. Daniel, VC, Peacock, CD, and Watkins, DN. Developmental signalling pathways in lung cancer. *Respirology* 2006;11: p. 234-40.
140. Olsen, CL, Hsu, PP, Glienke, J, Rubanyi, GM, and Brooks, AR. Hedgehog-interacting protein is highly expressed in endothelial cells but down-regulated during angiogenesis and in several human tumors. *BMC Cancer* 2004;4: p. 43.
141. Nickoloff, BJ, Osborne, BA, and Miele, L. Notch signaling as a therapeutic target in cancer: a new approach to the development of cell fate modifying agents. *Oncogene* 2003;22: p. 6598-608.
142. Rubin, LL and de Sauvage, FJ. Targeting the Hedgehog pathway in cancer. *Nat Rev Drug Discov* 2006;5: p. 1026-33.
143. Riobo, NA, Lu, K, and Emerson, CP, Jr. Hedgehog signal transduction: signal integration and cross talk in development and cancer. *Cell Cycle* 2006;5: p. 1612-5.
144. Lauth, M and Toftgard, R. Non-canonical activation of GLI transcription factors: implications for targeted anti-cancer therapy. *Cell Cycle* 2007;6: p. 2458-63.
145. Watkins, DN, Berman, DM, Burkholder, SG, Wang, B, Beachy, PA, and Baylin, SB. Hedgehog signalling within airway epithelial progenitors and in small-cell lung cancer. *Nature* 2003;422: p. 313-7.
146. Yuan, Z, Goetz, JA, Singh, S, et al. Frequent requirement of hedgehog signaling in non-small cell lung carcinoma. *Oncogene* 2007;26: p. 1046-55.

147. Dang, TP, Eichenberger, S, Gonzalez, A, Olson, S, and Carbone, DP. Constitutive activation of Notch3 inhibits terminal epithelial differentiation in lungs of transgenic mice. *Oncogene* 2003;22: p. 1988-97.
148. Politi, K, Feirt, N, and Kitajewski, J. Notch in mammary gland development and breast cancer. *Semin Cancer Biol* 2004;14: p. 341-7.
149. Parr, C, Watkins, G, and Jiang, WG. The possible correlation of Notch-1 and Notch-2 with clinical outcome and tumour clinicopathological parameters in human breast cancer. *Int J Mol Med* 2004;14: p. 779-86.
150. Hainaud, P, Contreres, JO, Villemain, A, et al. The role of the vascular endothelial growth factor-Delta-like 4 ligand/Notch4-ephrin B2 cascade in tumor vessel remodeling and endothelial cell functions. *Cancer Res* 2006;66: p. 8501-10.
151. Sansone, P, Storci, G, Tavolari, S, et al. IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland. *J Clin Invest* 2007;117: p. 3988-4002.
152. Haruki, N, Kawaguchi, KS, Eichenberger, S, et al. Dominant-negative Notch3 receptor inhibits mitogen-activated protein kinase pathway and the growth of human lung cancers. *Cancer Res* 2005;65: p. 3555-61.
153. Ishihara, K and Hirano, T. Molecular basis of the cell specificity of cytokine action. *Biochim Biophys Acta* 2002;1592: p. 281-96.
154. Hong, DS, Angelo, LS, and Kurzrock, R. Interleukin-6 and its receptor in cancer: implications for Translational Therapeutics. *Cancer* 2007;110: p. 1911-28.
155. Gao, SP, Mark, KG, Leslie, K, et al. Mutations in the EGFR kinase domain mediate STAT3 activation via IL-6 production in human lung adenocarcinomas. *J Clin Invest* 2007;117: p. 3846-56.
156. He, B, You, L, Uematsu, K, et al. A monoclonal antibody against Wnt-1 induces apoptosis in human cancer cells. *Neoplasia* 2004;6: p. 7-14.
157. You, L, He, B, Xu, Z, et al. Inhibition of Wnt-2-mediated signaling induces programmed cell death in non-small-cell lung cancer cells. *Oncogene* 2004;23: p. 6170-4.
158. Winn, RA, Van Scoyk, M, Hammond, M, et al. Antitumorigenic effect of Wnt 7a and Fzd 9 in non-small cell lung cancer cells is mediated through ERK-5-dependent activation of peroxisome proliferator-activated receptor gamma. *J Biol Chem* 2006;281: p. 26943-50.
159. Fukui, T, Kondo, M, Ito, G, et al. Transcriptional silencing of secreted frizzled related protein 1 (SFRP 1) by promoter hypermethylation in non-small-cell lung cancer. *Oncogene* 2005;24: p. 6323-7.
160. Mazieres, J, He, B, You, L, et al. Wnt inhibitory factor-1 is silenced by promoter hypermethylation in human lung cancer. *Cancer Res* 2004;64: p. 4717-20.
161. Uematsu, K, He, B, You, L, Xu, Z, McCormick, F, and Jablons, DM. Activation of the Wnt pathway in non small cell lung cancer: evidence of dishevelled overexpression. *Oncogene* 2003;22: p. 7218-21.
162. Wissmann, C, Wild, PJ, Kaiser, S, et al. WIF1, a component of the Wnt pathway, is down-regulated in prostate, breast, lung, and bladder cancer. *J Pathol* 2003;201: p. 204-12.
163. Syrjanen, KJ. Condylomatous changes in neoplastic bronchial epithelium. Report of a case. *Respiration* 1979;38: p. 299-304.

164. Klein, F, Kotb, WF, and Petersen, I. Incidence of human papilloma virus in lung cancer. *Lung Cancer* 2008.
165. Beer, DG, Kardia, SL, Huang, CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8: p. 816-24.
166. Bild, AH, Yao, G, Chang, JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439: p. 353-7.
167. Potti, A, Dressman, HK, Bild, A, et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006;12: p. 1294-300.
168. Shedden, K, Taylor, JM, Enkemann, SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14: p. 822-7.
169. Anguiano, A, Nevins, JR, and Potti, A. Toward the individualization of lung cancer therapy. *Cancer* 2008;113: p. 1760-7.
170. Xie, Y and Minna, JD. Predicting the future for people with lung cancer. *Nat Med* 2008;14: p. 812-3.
171. Tonon, G, Wong, KK, Maulik, G, et al. High-resolution genomic profiles of human lung cancer. *Proc Natl Acad Sci U S A* 2005;102: p. 9625-30.
172. Zhao, X, Weir, BA, LaFramboise, T, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. *Cancer Res* 2005;65: p. 5561-70.
173. Shibata, T, Uryu, S, Kokubu, A, et al. Genetic classification of lung adenocarcinoma based on array-based comparative genomic hybridization analysis: its association with clinicopathologic features. *Clin Cancer Res* 2005;11: p. 6177-85.
174. Ding, L, Getz, G, Wheeler, DA, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455: p. 1069-75.
175. Sekido, Y, Obata, Y, Ueda, R, et al. Preferential expression of c-kit protooncogene transcripts in small cell lung cancer. *Cancer Res* 1991;51: p. 2416-9.
176. Boldrini, L, Ursino, S, Gisfredi, S, et al. Expression and mutational status of c-kit in small-cell lung cancer: prognostic relevance. *Clin Cancer Res* 2004;10: p. 4101-8.
177. Shigematsu, H and Gazdar, AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int J Cancer* 2006;118: p. 257-62.
178. Yousem, SA, Nikiforova, M, and Nikiforov, Y. The histopathology of BRAF-V600E-mutated lung adenocarcinoma. *Am J Surg Pathol* 2008;32: p. 1317-21.
179. Hirsch, FR, Varella-Garcia, M, Bunn, PA, Jr., et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J Clin Oncol* 2003;21: p. 3798-807.
180. Nakamura, H, Saji, H, Ogata, A, et al. Correlation between encoded protein overexpression and copy number of the HER2 gene with survival in non-small cell lung cancer. *Int J Cancer* 2003;103: p. 61-6.
181. Hirashima, N, Takahashi, W, Yoshii, S, Yamane, T, and Ooi, A. Protein overexpression and gene amplification of c-erb B-2 in pulmonary carcinomas: a comparative immunohistochemical and fluorescence in situ hybridization study. *Mod Pathol* 2001;14: p. 556-62.
182. Tatematsu, A, Shimizu, J, Murakami, Y, et al. Epidermal growth factor receptor mutations in small cell lung cancer. *Clin Cancer Res* 2008;14: p. 6092-6.

183. Swanton, C, Futreal, A, and Eisen, T. Her2-targeted therapies in non-small cell lung cancer. *Clin Cancer Res* 2006;12: p. 4377s-4383s.

184. Rodenhuis, S, van de Wetering, ML, Mooi, WJ, Evers, SG, van Zandwijk, N, and Bos, JL. Mutational activation of the K-ras oncogene. A possible pathogenetic factor in adenocarcinoma of the lung. *N Engl J Med* 1987;317: p. 929-35.

185. De Biasi, F, Del Sal, G, and Hand, PH. Evidence of enhancement of the ras oncogene protein product (p21) in a spectrum of human tumors. *Int J Cancer* 1989;43: p. 431-5.

186. Potiron, VA, Roche, J, and Drabkin, HA. Semaphorins and their receptors in lung cancer. *Cancer Lett* 2008.

187. Samuels, Y, Wang, Z, Bardelli, A, et al. High frequency of mutations of the PIK3CA gene in human cancers. *Science* 2004;304: p. 554.

188. Lee, JW, Soung, YH, Kim, SY, et al. PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. *Oncogene* 2005;24: p. 1477-80.

189. Davies, H, Hunter, C, Smith, R, et al. Somatic mutations of the protein kinase gene family in human lung cancer. *Cancer Res* 2005;65: p. 7591-5.

190. Micke, P, Hengstler, JG, Ros, R, et al. c-erbB-2 expression in small-cell lung cancer is associated with poor prognosis. *Int J Cancer* 2001;92: p. 474-9.

191. Potti, A, Willardson, J, Forseen, C, et al. Predictive role of HER-2/neu overexpression and clinical features at initial presentation in patients with extensive stage small cell lung carcinoma. *Lung Cancer* 2002;36: p. 257-61.

192. Dworakowska, D, Jassem, E, Jassem, J, et al. MDM2 gene amplification: a new independent factor of adverse prognosis in non-small cell lung cancer (NSCLC). *Lung Cancer* 2004;43: p. 285-95.

193. Cappuzzo, F, Janne, PA, Skokan, M, et al. MET increased gene copy number and primary resistance to gefitinib therapy in non-small-cell lung cancer patients. *Ann Oncol* 2008.

194. Beau-Faller, M, Ruppert, AM, Voegeli, AC, et al. MET gene copy number in non-small cell lung cancer: molecular analysis in a targeted tyrosine kinase inhibitor naive cohort. *J Thorac Oncol* 2008;3: p. 331-9.

195. Richardson, GE and Johnson, BE. The biology of lung cancer. *Semin Oncol* 1993;20: p. 105-27.

196. Johnson, BE, Russell, E, Simmons, AM, et al. MYC family DNA amplification in 126 tumor cell lines from patients with small cell lung cancer. *J Cell Biochem Suppl* 1996;24: p. 210-7.

197. Ibson, JM, Waters, JJ, Twentyman, PR, Bleehen, NM, and Rabbitts, PH. Oncogene amplification and chromosomal abnormalities in small cell lung cancer. *J Cell Biochem* 1987;33: p. 267-88.

198. Shiraishi, M, Noguchi, M, Shimosato, Y, and Sekiya, T. Amplification of protooncogenes in surgical specimens of human lung carcinomas. *Cancer Res* 1989;49: p. 6474-9.

199. Miller, CT, Chen, G, Gharib, TG, et al. Increased C-CRK proto-oncogene expression is associated with an aggressive phenotype in lung adenocarcinomas. *Oncogene* 2003;22: p. 7950-7.

200. Pezzella, F, Turley, H, Kuzu, I, et al. bcl-2 protein in non-small-cell lung carcinoma. *N Engl J Med* 1993;329: p. 690-4.

201. Kaiser, U, Schilli, M, Haag, U, et al. Expression of bcl-2--protein in small cell lung cancer. *Lung Cancer* 1996;15: p. 31-40.
202. Reissmann, PT, Koga, H, Figlin, RA, Holmes, EC, and Slamon, DJ. Amplification and overexpression of the cyclin D1 and epidermal growth factor receptor genes in non-small-cell lung cancer. Lung Cancer Study Group. *J Cancer Res Clin Oncol* 1999;125: p. 61-70.
203. Eren, B, Sar, M, Oz, B, and Dincbas, FH. MMP-2, TIMP-2 and CD44v6 expression in non-small-cell lung carcinomas. *Ann Acad Med Singapore* 2008;37: p. 32-9.
204. Junker, K, Wiethege, T, and Muller, KM. Pathology of small-cell lung cancer. *J Cancer Res Clin Oncol* 2000;126: p. 361-8.
205. Micke, P, Basrai, M, Faldum, A, et al. Characterization of c-kit expression in small cell lung cancer: prognostic and therapeutic implications. *Clin Cancer Res* 2003;9: p. 188-94.
206. Cook, RM, Miller, YE, and Bunn, PA, Jr. Small cell lung cancer: etiology, biology, clinical features, staging, and treatment. *Curr Probl Cancer* 1993;17: p. 69-141.
207. Araki, K, Ishii, G, Yokose, T, et al. Frequent overexpression of the c-kit protein in large cell neuroendocrine carcinoma of the lung. *Lung Cancer* 2003;40: p. 173-80.
208. Rygaard, K, Nakamura, T, and Spang-Thomsen, M. Expression of the proto-oncogenes c-met and c-kit and their ligands, hepatocyte growth factor/scatter factor and stem cell factor, in SCLC cell lines and xenografts. *Br J Cancer* 1993;67: p. 37-46.
209. Plummer, H, 3rd, Catlett, J, Leftwich, J, et al. c-myc expression correlates with suppression of c-kit protooncogene expression in small cell lung cancer cell lines. *Cancer Res* 1993;53: p. 4337-42.
210. Hibi, K, Takahashi, T, Sekido, Y, et al. Coexpression of the stem cell factor and the c-kit genes in small-cell lung cancer. *Oncogene* 1991;6: p. 2291-6.
211. Weiner, DB, Nordberg, J, Robinson, R, et al. Expression of the neu gene-encoded protein (P185neu) in human non-small cell carcinomas of the lung. *Cancer Res* 1990;50: p. 421-5.
212. Schneider, PM, Hung, MC, Chiocca, SM, et al. Differential expression of the c-erbB-2 gene in human small cell and non-small cell lung cancer. *Cancer Res* 1989;49: p. 4968-71.
213. Fernandes, A, Hamburger, AW, and Gerwin, BI. ErbB-2 kinase is required for constitutive stat 3 activation in malignant human lung epithelial cells. *Int J Cancer* 1999;83: p. 564-70.
214. Rygaard, K, Vindelov, LL, and Spang-Thomsen, M. Expression of myc family oncoproteins in small-cell lung-cancer cell lines and xenografts. *Int J Cancer* 1993;54: p. 144-52.
215. Takahashi, T, Obata, Y, Sekido, Y, et al. Expression and amplification of myc gene family in small cell lung cancer and its relation to biological characteristics. *Cancer Res* 1989;49: p. 2683-8.
216. Spencer, CA and Groudine, M. Control of c-myc regulation in normal and neoplastic cells. *Adv Cancer Res* 1991;56: p. 1-48.
217. Zhang, P, Gao, WY, Turner, S, and Ducatman, BS. Gleevec (STI-571) inhibits lung cancer cell growth (A549) and potentiates the cisplatin effect in vitro. *Mol Cancer* 2003;2: p. 1.
218. Rikova, K, Guo, A, Zeng, Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 2007;131: p. 1190-203.

219. Johnson, FM, Krug, LM, Tran, HT, et al. Phase I studies of imatinib mesylate combined with cisplatin and irinotecan in patients with small cell lung carcinoma. *Cancer* 2006;106: p. 366-74.
220. Rossi, G, Cavazza, A, Marchioni, A, et al. Role of chemotherapy and the receptor tyrosine kinases KIT, PDGFRalpha, PDGFRbeta, and Met in large-cell neuroendocrine carcinoma of the lung. *J Clin Oncol* 2005;23: p. 8774-85.
221. Carbone, DP, Mitsudomi, T, Chiba, I, et al. p53 immunostaining positivity is associated with reduced survival and is imperfectly correlated with gene mutations in resected non-small cell lung cancer. A preliminary report of LCSG 871. *Chest* 1994;106: p. 377S-381S.
222. Wistuba, II, Berry, J, Behrens, C, et al. Molecular changes in the bronchial epithelium of patients with small cell lung cancer. *Clin Cancer Res* 2000;6: p. 2604-10.
223. Chiba, I, Takahashi, T, Nau, MM, et al. Mutations in the p53 gene are frequent in primary, resected non-small cell lung cancer. Lung Cancer Study Group. *Oncogene* 1990;5: p. 1603-10.
224. Shimizu, E, Zhao, M, Shinohara, A, et al. Differential expressions of cyclin A and the retinoblastoma gene product in histological subtypes of lung cancer cell lines. *J Cancer Res Clin Oncol* 1997;123: p. 533-8.
225. Salgia, R and Skarin, AT. Molecular abnormalities in lung cancer. *J Clin Oncol* 1998;16: p. 1207-17.
226. Hensel, CH, Hsieh, CL, Gazdar, AF, et al. Altered structure and expression of the human retinoblastoma susceptibility gene in small cell lung cancer. *Cancer Res* 1990;50: p. 3067-72.
227. Girard, L, Zochbauer-Muller, S, Virmani, AK, Gazdar, AF, and Minna, JD. Genome-wide allelotyping of lung cancer identifies new regions of allelic loss, differences between small cell lung cancer and non-small cell lung cancer, and loci clustering. *Cancer Res* 2000;60: p. 4894-906.
228. Virmani, AK, Fong, KM, Kodagoda, D, et al. Allelotyping demonstrates common and distinct patterns of chromosomal loss in human lung cancer types. *Genes Chromosomes Cancer* 1998;21: p. 308-19.
229. Thiberville, L, Payne, P, Vielkinds, J, et al. Evidence of cumulative gene losses with progression of premalignant epithelial lesions to carcinoma of the bronchus. *Cancer Res* 1995;55: p. 5133-9.
230. Sunaga, N, Miyajima, K, Suzuki, M, et al. Different roles for caveolin-1 in the development of non-small cell lung cancer versus small cell lung cancer. *Cancer Res* 2004;64: p. 4277-85.
231. Mori, S, Ito, G, Usami, N, et al. p53 apoptotic pathway molecules are frequently and simultaneously altered in nonsmall cell lung carcinoma. *Cancer* 2004;100: p. 1673-82.
232. Prudkin, L, Behrens, C, Liu, DD, et al. Loss and reduction of FUS1 protein expression is a frequent phenomenon in the pathogenesis of lung cancer. *Clin Cancer Res* 2008;14: p. 41-7.
233. Safar, AM, Spencer, H, 3rd, Su, X, et al. Methylation profiling of archived non-small cell lung cancer: a promising prognostic system. *Clin Cancer Res* 2005;11: p. 4400-5.
234. Toyooka, S, Toyooka, KO, Maruyama, R, et al. DNA methylation profiles of lung tumors. *Mol Cancer Ther* 2001;1: p. 61-7.

235. Shimamoto, T, Ohyashiki, JH, Hirano, T, Kato, H, and Ohyashiki, K. Hypermethylation of E-cadherin gene is frequent and independent of p16INK4A methylation in non-small cell lung cancer: potential prognostic implication. *Oncol Rep* 2004;12: p. 389-95.

236. Jarmalaite, S, Kannio, A, Anttila, S, Lazutka, JR, and Husgafvel-Pursiainen, K. Aberrant p16 promoter methylation in smokers and former smokers with nonsmall cell lung cancer. *Int J Cancer* 2003;106: p. 913-8.

237. Esteller, M, Sanchez-Cespedes, M, Rosell, R, Sidransky, D, Baylin, SB, and Herman, JG. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res* 1999;59: p. 67-70.

238. Kim, DH, Nelson, HH, Wiencke, JK, et al. p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. *Cancer Res* 2001;61: p. 3419-24.

239. Tang, X, Khuri, FR, Lee, JJ, et al. Hypermethylation of the death-associated protein (DAP) kinase promoter and aggressiveness in stage I non-small-cell lung cancer. *J Natl Cancer Inst* 2000;92: p. 1511-6.

240. Kim, DS, Cha, SI, Lee, JH, et al. Aberrant DNA methylation profiles of non-small cell lung cancers in a Korean population. *Lung Cancer* 2007;58: p. 1-6.

241. Suh, YA, Lee, HY, Virmani, A, et al. Loss of retinoic acid receptor beta gene expression is linked to aberrant histone H3 acetylation in lung cancer cell lines. *Cancer Res* 2002;62: p. 3945-9.

242. Burbee, DG, Forgacs, E, Zochbauer-Muller, S, et al. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *J Natl Cancer Inst* 2001;93: p. 691-9.

243. Dammann, R, Li, C, Yoon, JH, Chin, PL, Bates, S, and Pfeifer, GP. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. *Nat Genet* 2000;25: p. 315-9.

244. Speicher, MR, Gwyn Ballard, S, and Ward, DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 1996;12: p. 368-75.

245. Balsara, BR and Testa, JR. Chromosomal imbalances in human lung cancer. *Oncogene* 2002;21: p. 6877-83.

246. Luk, C, Tsao, MS, Bayani, J, Shepherd, F, and Squire, JA. Molecular cytogenetic analysis of non-small cell lung carcinoma by spectral karyotyping and comparative genomic hybridization. *Cancer Genet Cytogenet* 2001;125: p. 87-99.

247. Petersen, I, Bujard, M, Petersen, S, et al. Patterns of chromosomal imbalances in adenocarcinoma and squamous cell carcinoma of the lung. *Cancer Res* 1997;57: p. 2331-5.

248. Petersen, I, Langreck, H, Wolf, G, et al. Small-cell lung cancer is characterized by a high incidence of deletions on chromosomes 3p, 4q, 5q, 10q, 13q and 17p. *Br J Cancer* 1997;75: p. 79-86.

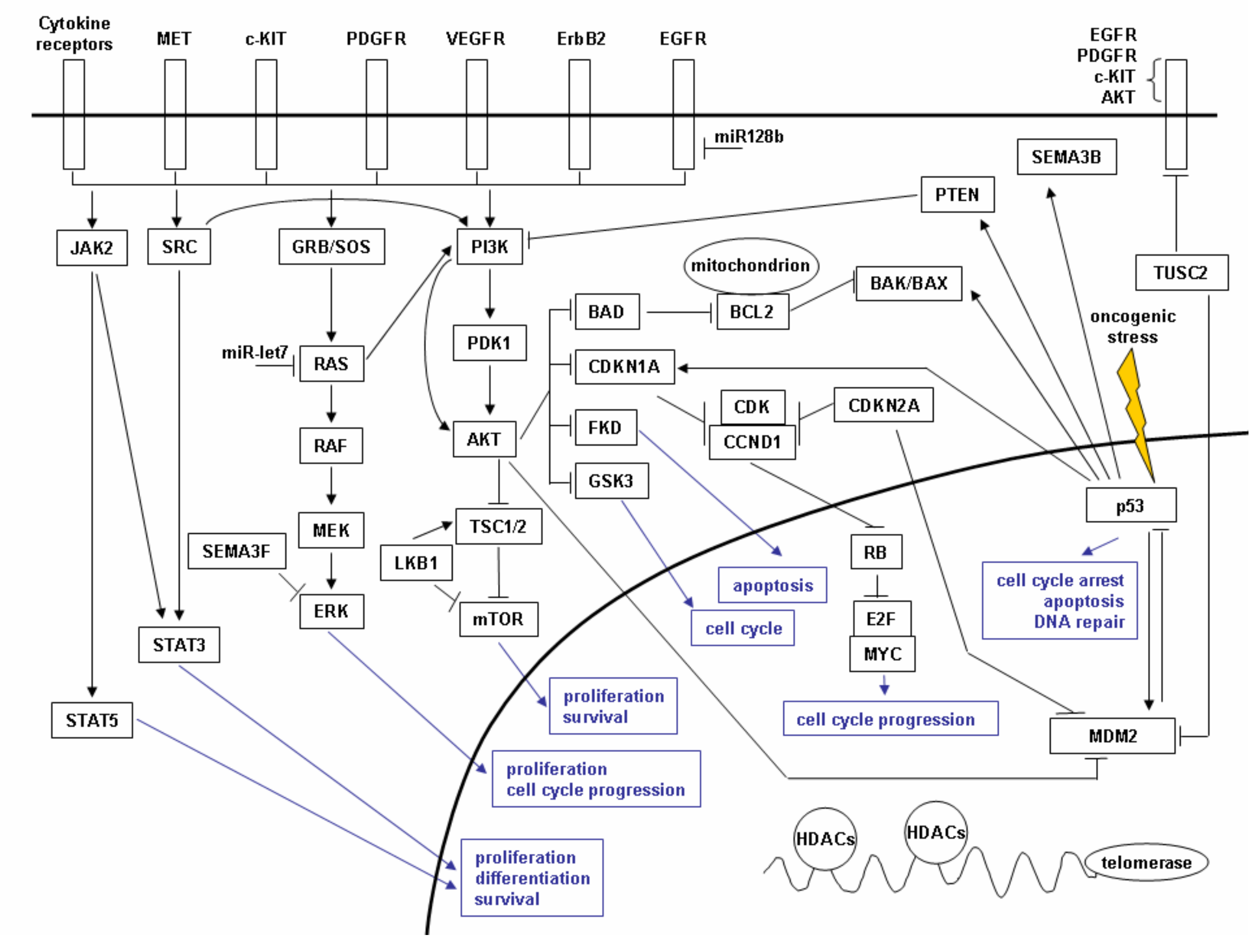

Cytokine receptors
MET
c-KIT
PDGFR
VEGFR
ErbB2
EGFR
EGFR
PDGFR
c-KIT
AKT
miR128b
SEMA3B
PTEN
JAK2
SRC
GRB/SOS
PI3K
mitochondrion
BAK/BAX
TUSC2
BAD
BCL2
oncogenic stress
PDK1
miR-let7
RAS
CDKN1A
CDK
CDKN2A
CCND1
RAF
AKT
FKD
GSK3
p53
TSC1/2
MEK
SEMA3F
LKB1
RB
apoptosis
cell cycle arrest
apoptosis
DNA repair
ERK
mTOR
E2F
cell cycle
STAT3
MYC
proliferation
survival
cell cycle progression
STAT5
MDM2
proliferation
cell cycle progression
HDACs
HDACs
telomerase
proliferation
differentiation
survival

Figure 1: Signaling pathways involved in NSCLC and SCLC. Aberrant signaling resulting in activation of growth stimulatory pathways or interference of growth inhibitory pathways has been implicated in lung cancer pathogenesis. Oncogenic activation typically occurs through amplification, mutation, rearrangement or over-expression. In lung cancer, commonly activated oncogenes include *EGFR*, *ERBB2*, *MYC*, *KRAS*, *MET*, *CCND1*, *CDK4* and *BCL2*. In contrast to oncogene activation, loss of TSG function is thought to require inactivation of both alleles – generally, LOH of one allele, and point mutation, epigenetic or transcriptional silencing will inactivate the second allele. In lung cancer, commonly inactivated TSGs include *TP53*, *RB1*, *CDKN2A*, *FHIT* and *PTEN*. Although EGFR plays a major role in lung cancer pathogenesis, several other receptor tyrosine kinases have been implicated such as members of the platelet-derived growth factor receptor (PDGFR) family c-KIT (expression is common in SCLC but rare in NSCLC (175, 176)) and MET (potentially associated with acquired resistance to EGFR TKIs). Table 2 lists targeted therapeutic agents which have been developed against many components of these signaling pathways.

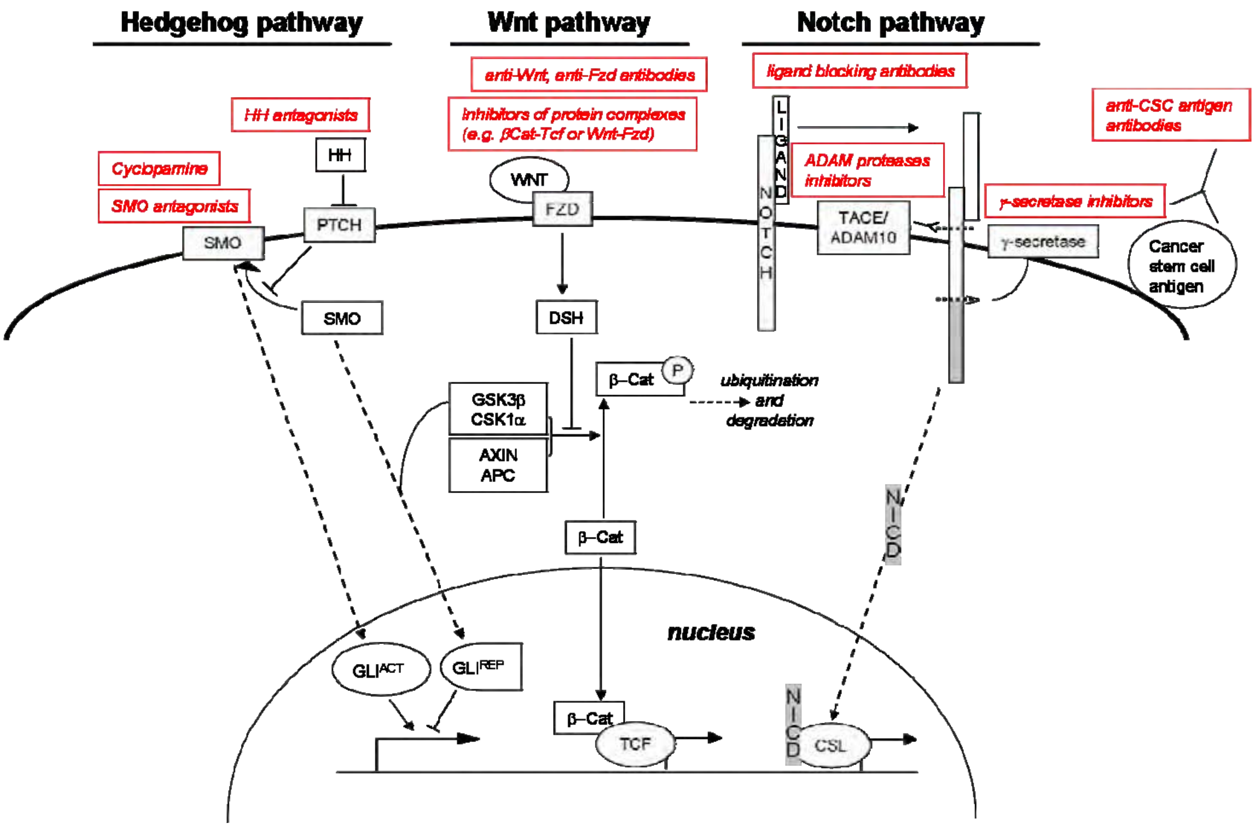

Hedgehog pathway
Wnt pathway
Notch pathway
anti-Wnt, anti-Fzd antibodies
ligand blocking antibodies
HH antagonists
inhibitors of protein complexes (e.g. βCat-Tcf or Wnt-Fzd)
anti-CSC antigen antibodies
Cyclopamine
SMO antagonists
HH
WNT
LIGAND
NOTCH
ADAM proteases inhibitors
γ-secretase inhibitors
SMO
PTCH
FZD
TACE/ ADAM10
γ-secretase
Cancer stem cell antigen
SMO
DSH
β-Cat
P
ubiquitination and degradation
GSK3β CSK1α
AXIN APC
β-Cat
NICD
nucleus
GLI^ACT
GLI^REP
β-Cat
TCF
NICD
CSL

Figure 2: Stem cell self-renewal pathways and therapeutic strategies to block these pathways in cancer. Notch, Wnt, and Hedgehog (HH) are stem cell self-renewal pathways that are often deregulated and aberrantly activated in lung cancer, thus representing key therapeutic targets. Modulation of these pathways can be achieved at different levels. In general, it is possible to interfere with ligand-receptor interactions by using ligand antagonists or receptor decoys, by blocking ligand-induced conformational changes, or disrupting protein-protein interaction of complexes involved in the activation of nuclear transcription. More specifically, strategies targeting the HH pathway include the use of antagonists for ligands (HH) or receptors (Smoothened, SMO) as well as cyclopamine, a naturally occurring compound that inhibits SMO. Antibodies against WNT ligands or the receptor FZD and inhibitors of protein complex formation can be used to inhibit the Wnt pathway. Anti-ligand antibodies are used as potential agents to block the Notch pathway. Cleavage of Notch receptors by ADAM proteases and γ-secretase is required to activate this pathway and inhibitors of these enzymes are being tested for their possible therapeutic implications. At present, two drugs targeting these pathways are in early clinical trials for treatment of some cancer types: phase I trials are studying the effects of a HH antagonist (GDC-0449) on solid tumors and MK-0752, a γ-secretase inhibitor, is being tested in patients with T-cell acute lymphoblastic leukemia, breast cancer, and central nervous system tumors. Pathways are depicted schematically and some components were omitted (dashed lines) for simplicity. ADAM, a disintegrin and metalloprotease domain; APC, adenomatous polyposis coli; β-Cat, β-catenin; CSK1α, cyclin suppressing kinase 1α; CSL, C-promoter binding factor 1, suppressor of hairless, and Lag1 protein complex; DSH, disheveled; FZD, frizzled; GLI, glioma-associated oncogene; GLI^{ACT}, active form of GLI; GLI^{REP}, repressor form of GLI; GSK3β, glycogen synthase kinase 3β; HH, hedgehog; NICD, Notch intracellular domain; P, phosphorylation; PTCH, patched; SMO, smoothened; TACE, TNF-α-converting enzyme; TCF/HNF4A, transcription factor.

Table 1: Common genetic alterations found in lung cancer [a]

Gene	SCLC (%)	NSCLC (%)			References
		All	Adeno-carcinoma	Squamous cell	
Oncogenic alterations					
Mutation					
BRAF	Rare	1-3	1-5	3	(177, 178)
EGFR	Rare	~20	10-40	Rare	(177, 179-182)
ErbB2 (HER2)	Rare	2	4	Rare	(177, 183)
KRAS	Rare	10-30	15-35	<5	(177, 184-186)
MET	13	21	14	12	(6)
PIK3CA	Rare	1-5	2-3	2-7	(56, 187-189)
Amplification					
EGFR	Rare	20-30	15	30	(6)
ErbB2 (HER2)	5-30	2-23	6	2	(6, 183, 190, 191)
MDM2		6-24	14	22	(192)
MET		7-21	20	21	(193, 194)
MYC	18-30	8-22			(195-198)
NKX2-1 (TITF1)	Rare	12-30	10-15	3-15	(6, 14, 15)
PIK3CA	~5	9-17	6	33-36	(6, 56)
Increase in protein expression					
CRK		8-30	8-30		(199)
BCL2	75-95	10-35			(186, 200, 201)
CCND1	0	43	35-55	30-35	(85, 202)
CD44	Rare	Common	3	48	(203)
c-KIT	46-91	Rare			(204-210)
EGFR	Rare	50-90	40-65	60-85	(29-32, 186)
ErbB2 (HER2)	<10	20-35	16-38	6-16	(183, 186, 207, 211-213)
MYC	10-45	<10			(50, 214-216)
PDGFRA	65	2-100	100	89	(217-220)
Tumor suppressing alterations					
Mutation					
CDKN2A (p16)	<1	10-40			(186)
LKB1	Rare	30	34	19	(6, 186)
p53	75-90	50-60	50-70	60-70	(186, 221-223)
PTEN	15-20	<10			(186)
Rb	80-100	20-40			(186, 224-226)
Deletion/LOH [b]					
CDKN2A (p16)	37	75-80			(84, 227, 228)
FHIT	100	55-75			(227-229)
p53	86-93	74-86			(227, 228)
Rb	93	62			(227, 228)
Loss of protein expression					
CAV1	95	24			(230)
CDKN2A ($p14^{ARF}$)	65	40-50			(84, 231)
CDKN2A (p16)	3-37	30-79	~55	60-75	(227, 228)
FHIT	80-95	40-70			(186, 227, 228)
PTEN		25-74	77	70	(60, 231)

Rb	90	15-60	23-57	6-14	(228)
TUSC2 (FUS1)	100	82	79	87	(232)
Tumor-acquired DNA Methylation					
APC	15-26	24-96			(93, 94, 233)
CAV1	93	9			(230)
CDH1	60 40	20-35			(94, 233-235)
CDH13	15-20	45			(93, 94)
CDKN2A (p14^{ARF})	nd[b]	6-8			(94)
CDKN2A (p16)	5, 0	15-41	21-36	24-33	(236-238)
DAPK1	nd	16-45			(94, 233, 239)
FHIT	64	37			(93, 94)
GSTP1	16	7-12, 15			(94, 240)
MGMT	16	16-27, 10			(94, 233)
PTEN		26	24	30	(60)
RARβ	45-70	40-43			(93, 94, 241)
RASSF1A	72-85	15-45	31	43	(90, 94, 96, 233, 242, 243)
SEMA3B	nd	41-50	46	47	(90, 91)
TIMP3	nd	19-26			(94)
Telomeres					
Telomerase activity	75-100	50-80	65-85	80-90	(11-13, 186, 244)
Chromosomal aberrations					
Large-scale loss	1p, 3p, 4p, 4q, 5q, 8p, 10q, 13q, 17p	3p, 5q, 8p, 9p, 13q, 17p, 18q, 19p, 19q, 21q, 22q	2q, 3p, 4q, 8p, 9p, 9q, 10p, 10q, 13q, 15q, 18, 20	3p, 4q, 9p, 10p, 10q, 18, 20	(24, 25, 55, 227, 245-248)
Focal deletions		2q22.1, 3p14.2, 3q25.1, 5q11.2, 7q11.22, 7q34, 9p23, 9p21.3, 10q23.31, 11q11, 13q12.11, 13q14.2, 13q32.2, 18q23, 21p11.2			(15, 171, 172)
Large-scale gain	3q, 5p, 8q, 18q	1q, 3q, 5p, 6p, 7p, 7q, 8q, 20p, 20q	5p, 7p, 7q, 8q, 11q, 19, 20q	2q, 3q, 5p, 7, 8q, 11q, 13q, 19, 20q	(24, 25, 55, 227, 245-248)
Focal amplifications		1p36.32, 1p34.3, 1q32.2, 1q21.2, 2p24.3, 2q11.2, 2q31.1, 3q26.31, 5p15.33, 5p15.31, 5p14.3, 5q31.3, 6p21.1, 7p11.2, 8p12, 8q21.13, 8q24.21, 10q24.1, 10q26.3, 11q13.3, 12p12.1, 12q13.2, 12q14.1, 12q15, 14q13.3, 14q32.13, 16q22.2, 17q12, 18q12.1, 19q12, 19q13.33, 20q13.32, 22q11.21			(15, 171, 172)

[a] nd, not determined

[b] LOH, loss of heterozygosity

Table 2: Targeted therapies against oncogenic pathways in lung cancer and their development in clinical trials

Gene	Drug	NSCLC[a]	SCLC[a]
EGFR	cetuximab	II/III	I[b]
EGFR	panitumumab	II	nct
EGFR	matuzumab	II	nct
EGFR	gefitinib	approved	II
EGFR	erlotinib	approved	nct
EGFR, ErbB2	lapatinib	II	nct
EGFR, ErbB2	HKI-272	II	nct
EGFR, ErbB2	CI-1033	II	nct
ErbB2	trastuzumab	II	I[b]
VEGF	bevacizumab	approved	II
VEGFR	cediranib	II/III	II
VEGFR, EGFR	vandetanib	II/III	II
VEGFR, PDGFR, c-KIT	sunitinib	II/III	II
VEGFR, PDGFR, c-KIT	vatalanib	II	nct
VEGFR, PDGFR, c-KIT	axitinib	II	nct
VEGFR, PDGFR, c-KIT	AMG-706	II/III	nct
c-KIT, PDGFR	Imatinib	II	II
RAS	tipifarnib	II	II
RAS	lonafarnib	III	nct
RAF, VEGFR, PDGFR, c-KIT	sorafenib	II/III	II
MEK	CI-1040	II	nct
MEK	PD-325901	II	nct
MEK	AZD6244	II	nct
PI3K	LY294002	nct	nct
mTOR	sirolimus	I/II	nct
mTOR	temsirolimus	II	II
mTOR	everolimus	I/II	I/II
mTOR	AP23573/deforolimus	I[b]	I[b]
BCL2	oblimersen	II/III	I/II
BCL2	ABT-737	nct	nct
SRC	dasatinib	II	II
Proteasome	bortezomib	II	II
Proteasome	NPI-0052	I	nct
p53	p53 peptide vaccine	II	I[b]
FUS1	fus1 liposome complex	I	nct
HDACs	vorinostat	II	I/II
HDACs	romidepsin	II	II
Telomerase	GRN163L	I	I[b]

[a] I, Phase I clinical trial; II, Phase II clinical trial; III, Phase III clinical trial; approved, approved by the FDA; nct, not in a clinical trial at time of manuscript preparation

[b] In Phase I clinical trial for solid tumors, not specific to NSCLC/SCLC

JNCI Journal of the National Cancer Institute Advance Access published September 9, 2008

BRIEF COMMUNICATION

Familial Aggregation of Common Sequence Variants on 15q24-25.1 in Lung Cancer

Pengyuan Liu, Haris G. Vikis, Daolong Wang, Yan Lu, Yian Wang, Ann G. Schwartz, Susan M. Pinney, Ping Yang, Mariza de Andrade, Gloria M. Petersen, Jonathan S. Wiest, Pamela R. Fain, Adi Gazdar, Colette Gaba, Henry Rothschild, Diptasri Mandal, Teresa Coons, Juwon Lee, Elena Kupert, Daniela Seminara, John Minna, Joan E. Bailey-Wilson, Xifeng Wu, Margaret R. Spitz, Timothy Eisen, Richard S. Houlston, Christopher I. Amos, Marshall W. Anderson, Ming You

Three recent genome-wide association studies identified associations between markers in the chromosomal region 15q24-25.1 and the risk of lung cancer. We conducted a genome-wide association analysis to investigate associations between single-nucleotide polymorphisms (SNPs) and the risk of lung cancer, in which we used blood DNA from 194 case patients with familial lung cancer and 219 cancer-free control subjects. We identified associations between common sequence variants at 15q24-25.1 (that spanned LOC123688 [a hypothetical gene], *PSMA4, CHRNA3, CHRNA5,* and *CHRNB4)* and lung cancer. The risk of lung cancer was more than five-fold higher among those subjects who had both a family history of lung cancer and two copies of high-risk alleles rs8034191 (odds ratio [OR] = 7.20, 95% confidence interval [CI] = 2.21 to 23.37) or rs1051730 (OR = 5.67, CI = 2.21 to 14.60, both of which were located in the 15q24-25.1 locus, than among control subjects. Thus, further research to elucidate causal variants in the 15q24-25.1 locus that are associated with lung cancer is warranted.

J Natl Cancer Inst 2008;100:1326–1330

Lung cancer can occur sporadically in people with no known family history of lung cancer or it can occur in multiple members of the same family and be designated as familial lung cancer. Evidence of a genetic basis for susceptibility to lung cancer has been demonstrated through genome-wide association studies (1–3) and segregation analyses (4–9). We conducted a genome-wide association study among individuals with a familial history of lung cancer. These individuals are members of families with three or more first-degree relatives with lung cancer that were collected as part of the Genetic Epidemiology of Lung Cancer Consortium (GELCC).

For the genome-wide association study, we genotyped 194 case patients with familial lung cancer and 219 cancer-free control subjects by use of the Affymetrix 500K or Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA) (Supplementary Table 1, available online). To ensure genetic independence among subjects, one case patient with familial lung cancer was chosen from each high-risk lung cancer family. Noncancer control subjects were obtained from a combination of unaffected spouses from GELCC families (n = 36) and of unaffected individuals from the Coriell Institute for Medical Research (Camden, NJ) (n = 11) and the Fernald Medical Monitoring Program (Fernald, OH) (n = 172). These control subjects had no blood relationship with any selected case patients. To minimize possible effects of cigarette smoking and age, smokers with an older age were selected mainly as control subjects, except for spousal control subjects (Supplementary Table 1, available online). To maintain homogeneous population samples, only Caucasian subjects from the GELCC, Coriell Institute for Medical Research, and the Fernald Medical Monitoring Program were used for the association analysis. Basic characteristics of the GELCC subjects are presented in Supplementary Table 1 (available online). The Texas and UK lung cancer cohorts have been described previously (1). Written informed consent and institutional review board approval were obtained for all the subjects involved in this study.

Single-nucleotide polymorphism (SNP) genotyping was done by the Vanderbilt University Microarray Shared Resource by following the Affymetrix protocol (www.affymetrix.com). We used the Affymetrix 500K chipset to genotype 76 control subjects and 73 case patients and the Affymetrix SNP 6.0 array to genotype the remaining 137 case patients and 164 control subjects. The samples used with the two platforms were randomly chosen from the GELCC collections. The Affymetrix 6.0 array includes more than 906 600 SNPs and covers 97.2% of SNPs presented in the Affymetrix 500K chipset of 500 568 SNPs. For the Affymetrix 500K SNP chipset, a confidence score of 0.33 was used for genotype determination, by using the Bayesian robust linear model with the Mahalanobis algorithm (10); an average genotyping call rate of 96.9% was obtained across all case and control samples. For the SNP Array 6.0, a block size of zero and a confidence threshold of 0.1 were used for genotype

---

Affiliations of authors: Washington University, St Louis, MO (PL, HGV, DW, YL, YW, MY); Karmanos Cancer Institute, Detroit, MI (AGS); University of Cincinnati, Cincinnati, OH (SMP, JLM, WA); Mayo Clinic, Rochester, MN (PY, MdA, GMP, DS); National Cancer Institute, Bethesda, MD (JSW); University of Colorado, Denver, CO (PRF); University of Texas Southwestern Medical Center, Dallas, TX (AG, JM); University of Toledo College of Medicine, Toledo, OH (CG); Louisiana State University Health Science Center, New Orleans, LA (HR, DM); Saccomanno Research Institute, Grand Junction, CO (TC); National Human Genome Research Institute, Baltimore, MD (JEB-W); M. D. Anderson Cancer Center, Houston, TX (XW, MRS, CIA); Department of Oncology, University of Cambridge, Cambridge, CB2 2RE, UK (TE); Section of Cancer Genetics, Institute of Cancer Research, SM2 5NG, UK (RSH).

Correspondence to: Ming You, MD, PhD, Department of Surgery and The Alvin J. Siteman Cancer Center, Washington University, 660 Euclid Ave, Box 8109, St Louis, MO 63110 (e-mail: youm@wudosis.wustl.edu).

See “Funding” and “Notes” following “References.”

DOI: 10.1093/jnci/djn268

© The Author 2008. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.

CONTEXT AND CAVEATS

Prior knowledge
Genome-wide association studies have identified associations between markers in the chromosomal region 15q24-25.1 and the risk of lung cancer.

Study design
A genome-wide case–control association analysis was used to investigate relationships between single-nucleotide polymorphisms (SNPs) and the risk of familial lung cancer.

Contribution
Subjects with both a family history of lung cancer and two copies of either of two high-risk alleles in 15q24-25.1 had a higher risk of lung cancer than control subjects.

Implications
Additional research is required to identify which genetic variants in the 15q24-25.1 region are associated with a high risk of lung cancer.

Limitations
Associations of risk alleles with nicotine dependence were not directly tested because the data were not available. Smoking quantity was available; however, no association between smoking quantity and the high-risk alleles was found. The small sample size may have limited the ability to detect a smaller effect size for risk alleles among heterozygotes with familial lung cancer.

From the Editors

---

determination, by using the Birdseed algorithm (available at http://www.broad.mit.edu/mpg/birdsuite/birdseed.html). We excluded four samples with unexpected genetic relatedness that was detected by PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/) and 33 samples with a genotype call rate of less than 86%. Thus, a total of 413 samples were available for our analysis (Supplementary Table 1, available online)—57 case patients and 56 control subjects genotyped with the Affymetrix 500K chipset, and 137 case patients and 163 control subjects genotyped with the Affymetrix SNP 6.0 array. When the frequency distribution of the overlapping SNPs in the two genotyping platforms was analyzed, no heterogeneity between the platforms was found (Supplementary Figure 1, available online). Among the 413 subjects, we detected no population stratification, as shown by linkage agglomerative clustering implemented in PLINK software. Hardy–Weinberg equilibrium for each SNP was examined with the software hweStrata, by use of an exact test as proposed previously (11) (stratum number $K = 1$). All statistical tests were two-sided. SNPs with an exact P value of .0001 or less from Hardy–Weinberg equilibrium tests and with a minor allele frequency of less than 1% among control subjects were excluded from association analyses. Thus, in the final association analysis of 413 samples, we retained 722 376 SNPs for the 300 samples that were genotyped by SNP Array 6.0 and 399 377 SNPs for the 113 samples that were genotyped by Affymetrix 500K. The statistical significance of the association between SNP allele and disease status was assessed primarily with Fisher exact tests (12) (Supplementary Figure 2, available online). The odds ratios (ORs) of lung cancer associated with each SNP and 95% confidence intervals (CIs) were estimated by allele and by genotype. We evaluated the cumulative distribution of P values from Fisher exact tests in our genome-wide association study. As shown in Supplementary Figure 3 (available online), the distribution of observed P values was similar to the expected uniform distribution [0, 1], indicating no inflation of test statistics from population structure or any other form of bias.

Our genome-wide association study identified several strong associations of SNPs on chromosomes 1, 3, 6, 9, 12, 15, and 20 with lung cancer. One of the most strongly associated clusters of SNPs was found in a 160-kb region at 15q24-25.1 (Figure 1, A and B, and Supplementary Figure 1, available online) that had strong linkage disequilibrium and contained multiple discrete haplotype blocks (ie, regions with no evidence of historical recombination), as defined by examining the HapMap data from Utah residents with ancestry from northern and western Europe (http://www.hapmap.org/) (Figure 1, C). These familial data confirm the recently described association between 15q25.1 and sporadic lung cancer (1–3). At least two common variants on 15q24-25.1, SNPs rs8034191 and rs1051730, were strongly associated with both familial ($P = 3.74 \times 10^{-4}$ for rs8034191 and $P = 2.90 \times 10^{-4}$ for rs1051730) and sporadic (from the Texas study, $P = 9.71 \times 10^{-8}$ for rs8034191 and $P = 1.32 \times 10^{-7}$ for rs1051730; from the UK study, $P = 9.90 \times 10^{-9}$ for rs8034191 and $P = 2.84 \times 10^{-8}$ for rs1051730) lung cancer (Table 1). These SNPs are in high linkage disequilibrium ($r^2 = 0.87$) and are located within intronic regions of hypothetical gene LOC123688 (rs8034191) and of the gene *CHRNA3* (rs1051730). Risk of lung cancer among homozygous carriers was statistically significantly different from risk among noncarriers for the two SNPs—rs8034191 and rs1051730—in large sporadic lung cancer samples previously reported from Texas and from the United Kingdom (1).

It is worth noting that the effect size of the risk allele at 15q24-q25.1 observed in familial lung cancer (for carrying two copies of risk alleles of rs8034191 and rs1051730, respectively, OR = 3.84, 95% CI = 1.75 to 8.84, and OR = 3.43, 95% CI = 1.66 to 7.37) was larger than that observed in sporadic lung cancer (for carrying two copies of the same risk alleles, respectively, OR = 1.76, 95% CI = 1.42 to 2.18, and OR = 1.75, 95% CI = 1.42 to 2.17, in the Texas study; and OR = 2.09, 95% CI = 1.60 to 2.73, and OR = 2.05, 95% CI = 1.57 to 2.70, in the UK study) (Table 1). We therefore tested heterogeneity between familial and sporadic lung cancers by Woolf's test (13). Statistically significant heterogeneity was observed between familial and the Texas sporadic lung cancer samples ($P = .04$) but not between familial and UK sporadic lung cancer samples ($P = .12$).

To rule out confounding effects of smoking behavior on lung cancer risk, the association analysis was adjusted by sex, age, and pack-years of cigarette exposure, as continuous variables (Supplementary Table 2, available online). Association of the 15q24-25.1 locus with familial lung cancer remained statistically significant ($P = 1.03 \times 10^{-3}$ for rs8034191 and $P = 3.10 \times 10^{-4}$ for rs1051730) after this adjustment. The adjusted associations were even stronger among those in lung cancer families carrying two copies of risk alleles of rs8034191 (OR = 7.20, 95% CI = 2.21 to 23.37) or rs1051730 (OR = 5.67, 95% CI = 2.21 to 14.60). Thus, among lung cancer patients with a family history, alleles associated with a high risk of lung cancer appear to be located on the q arm of chromosome 15.

We next investigated whether the SNPs on 15q24-q25.1 that were strongly associated with lung cancer acted additively, recessively, or dominantly by use of logistic regression and Bayesian information criteria. Given any two estimated models, the model with the lower value of Bayesian information criteria is the one to be preferred. The strongest associations were found with the recessive model (Table 1 and Supplementary Table 3, available online). In the recessive model, an increased risk of lung cancer was associated with having two copies of the high-risk A allele of rs1051730, compared with having no copies (OR = 3.43, 95% CI = 1.66 to 7.37, *P* value = 2.90×10^{-4}). However, an additive model gave similar levels of statistical significance for many of the most statistically significant SNPs in this region. The homozygous risk genotype appeared to occur more frequently and to have larger effect size among familial lung cancer samples (OR = 3.84 for rs1051730) than among sporadic lung cancer samples (OR = 1.75 for rs1051730 from the Texas study and OR = 2.09 for rs1051730 in the UK study). The smaller sample size of the familial study was less powerful to detect the smaller risks among heterozygotes. This limitation and the higher frequency of homozygotes in the highly ascertained families with familial lung cancer may account for the difference in best fitting models at the 15q24 locus in familial vs sporadic data sets. A genome-wide association study on a larger sample of familial case patients would provide increased power. However, both familial and sporadic analyses support the presence of a risk allele at 15q24-25.1.

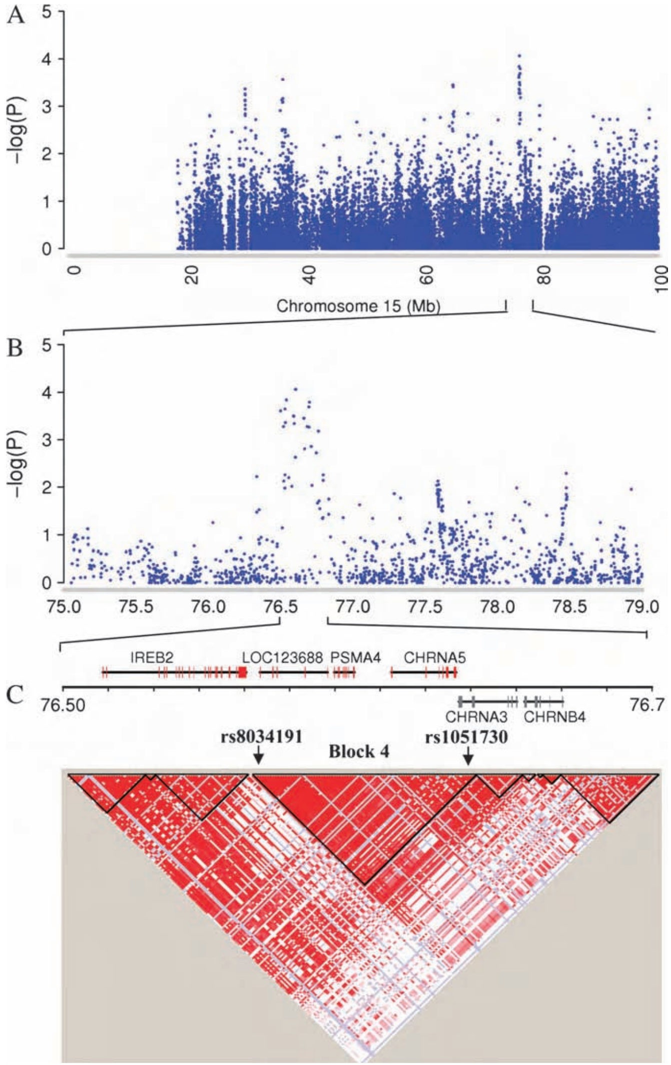


Figure 1. Association between chromosome 15q24-25.1 and lung cancer. Associations are expressed as −log(*P*); *P* values were from Fisher exact tests. All statistical tests were two-sided. **A**) Association analysis of 194 familial case patients and 219 control subjects in chromosome 15. **B**) Enhanced view of association analysis in the 75- to 79-Mb region of chromosome 15. Several common variants, including rs8034191 and rs1051730, cluster on 15q24-25.1 and show strong associations with familial lung cancer. **C**) Physical map and haplotype blocks (ie, regions with no evidence of historical recombination). Pairwise linkage disequilibrium, measured as *D*′ (a measurement of the nonrandom association of alleles at two loci), was calculated with HapMap data from Utah residents with ancestry from northern and western Europe by the methods of Gabriel (21) as implemented with Haploview software (22). Shading represents the magnitude and statistical significance of the pairwise linkage disequilibrium, with a white-to-red gradient reflecting lower to higher linkage disequilibrium values. The single-nucleotide polymorphisms rs8034191 and rs1051730 were contained in the 84-kb block 4.

In addition to analyses of single SNPs, we performed a haplotype-based association analysis on the 15q24-q25.1 locus. Haplotypes at the 15q24-25.1 locus region were inferred by use of the fastPHASE computer program (14). To exhaustively exploit haplotype information, we then subject alleles (contiguous sets of markers) from sliding windows of all sizes to haplotype association tests (15). In the haplotype analysis, a common haplotype with a frequency larger than 5% was first identified and treated as one category and the other variants as another category. The Fisher exact test was then applied to determine statistical significance (12). Multiple haplotypes at the 15q24-25.1 locus were statistically significantly associated with familial lung cancer risk (Supplementary Table 4, available online). The most statistically significant association was with the haplotype A-T consisting of both rs7163730 and rs4461039 (*P* = 1.03×10^{-4}). In the single-marker analyses, the two SNPs with the most statistically significant association with lung cancer were rs7163730 (*P* = 8.71×10^{-5}) and rs4461039 (*P* = 8.67×10^{-5}).

Using GELCC familial lung cancer samples and a genome-wide association study, we identified several common variants at 15q24-25.1 and confirmed the association between genetic variants on the q arm of chromosome 15 and sporadic lung cancer (1–3). When we adjusted for smoking and other covariates, a statistically significant association remained (Supplementary Table 2, available online). This observation is consistent with previous analyses (1–3), in which a direct role for these variants in lung cancer was postulated. In those analyses, the

Table 1. Associations of the single-nucleotide polymorphisms rs8034191 and rs1051730 genotypes with lung cancer*

		Frequency, %			
Samples†	**Genotype**	**Control subjects**	**Case patients**	***P* value‡**	**OR (95% CI)**
GELCC					
rs8034191	AA	50.00	36.64		
	AG	41.67	39.69	.362	1.30 (0.75 to 2.24)
	GG	8.33	23.66	3.74×10^{-4}	3.84 (1.75 to 8.84)
rs1051730§	GG	48.11	37.57		
	AG	44.81	43.35	.377	1.23 (0.78 to 1.96)
	AA	7.08	19.08	2.90×10^{-4}	3.43 (1.66 to 7.37)
Texas					
rs8034191	AA	44.10	37.01		
	AG	44.85	46.68	3.09×10^{-3}	1.24 (1.07 to 1.43)
	GG	11.05	16.32	9.71×10^{-8}	1.76 (1.42 to 2.18)
rs1051730	GG	44.31	36.82		
	AG	44.54	46.95	1.04×10^{-3}	1.27 (1.10 to 1.47)
	AA	11.15	16.23	1.32×10^{-7}	1.75 (1.42 to 2.17)
United Kingdom					
rs8034191	AA	46.67	36.59		
	AG	43.23	46.86	1.72×10^{-4}	1.38 (1.17 to 1.64)
	GG	10.10	16.55	9.90×10^{-9}	2.09 (1.60 to 2.73)
rs1051730	GG	46.55	37.54		
	AG	43.72	46.34	1.40×10^{-3}	1.31 (1.11 to 1.56)
	AA	9.73	16.12	2.84×10^{-8}	2.05 (1.57 to 2.70)

* GELCC = Genetic Epidemiology of Lung Cancer Consortium; OR = odds ratio; CI = confidence interval.

† The GELCC study had a total of 194 familial case patients and 219 control subjects. The Texas study had a total of 1865 sporadic case patients and 1769 control subjects. The UK study had a total of 2013 sporadic case patients and 3062 control subjects. The Texas and UK sporadic lung cancer samples are from the study by Amos et al. (1).

‡ Fisher exact tests were performed. All statistical tests were two-sided.

§ rs1051730 was genotyped by Human Genetics Division Genotyping Core.

number of cigarettes consumed per day, pack-years of exposure, and time to first cigarette were not associated with SNPs in the region (1,2). In contrast, Thorgeirsson et al. (3) reported that the SNP rs1051730 was statistically significantly associated with lung cancer and that each copy of the T allele was associated with an increase in smoking of one cigarette per day. In addition, Saccone et al. (16) identified the SNPs rs16969968 (missense in *CHRNA5*) and rs578776 (in the 3′-untranslated region of *CHRNA3*) as functional variants that were statistically significantly associated with nicotine dependence. When we genotyped rs16969968 and rs578776 in our familial lung cancer population, we found that both were statistically significantly associated with lung cancer (for rs16969968, $P = 2.29 \times 10^{-3}$; for rs578776, $P = 4.47 \times 10^{-4}$). Thus, polymorphisms in these genes may affect nicotine dependence and propensity to smoke and to develop lung cancer.

Our study had several limitations. First, the association of risk variants with nicotine dependence (as measured by the Fagerstrom test for nicotine dependence) was not directly tested because these data were unavailable. However, smoking quantity (a component of the Fagerstrom test) was available. We did not find an association between smoking quantity and the statistically significant variants associated with lung cancer. Second, the sample size in our familial genome-wide association analysis was small, and so we might not have been able to detect the smaller effect size of risk alleles located in 15q24-25.1 for heterozygotes in familial lung cancer.

The identity of variants in the 15q24-25.1 region that are most strongly associated with lung cancer remains unknown because many SNPs are in strong linkage disequilibrium with each other. In addition, whether the variants associate with a direct or indirect mode for lung cancer remains unresolved. From results of this and three other genome-wide association studies, the candidate genes include *IREB2*, LOC123688, *PSMA4*, *CHRNA5*, *CHRNA3*, and *CHRNB4* (1–3). *CHRNA5*, *CHRNA3*, and *CHRNB4* are strong candidates primarily because they encode subunits of the nicotinic acetylcholine receptors. These genes may participate in nicotinic addiction through reward pathways in the brain; however, evidence exists that nicotinic acetylcholine receptors are associated directly with lung carcinogenesis (17). Nicotinic acetylcholine receptors are expressed in lung cancers of both non–small cell and small cell subtypes (18), and treatment of lung cancer cell lines with nicotine can inhibit proapoptotic pathways initiated by opioids (19,20). Therefore, determination of a likely single candidate gene and further delineation of whether variants affect lung cancer directly or indirectly or both are warranted.

References

1. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet.* 2008;40(5):616–622.
2. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* 2008;452(7187):633–637.
3. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008;452(7187):638–642.

4. Sellers TA, Bailey-Wilson JE, Elston RC, et al. Evidence for Mendelian inheritance in the pathogenesis of lung cancer. *J Natl Cancer Inst.* 1990;82(15):1272–1279.
5. Bailey-Wilson JE, Elston RC, Sellers TA, Rothschild H. Segregation analysis of lung cancer using class A regressive models. *Am J Hum Genet.* 1992;51:A145.
6. Sellers TA, Bailey-Wilson JE, Elston RC, Rothschild H. Evidence for Mendelian factors in early onset lung cancer. In: Brugge J, Carran T, Harlow E, McCormick F, eds. *Origins of Human Cancer: A Comprehensive Review*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1991:775–780.
7. Sellers TA, Potter JD, Folsom AR. Association of incident lung cancer with family history of female reproductive cancers: the Iowa Women's Health Study. *Genet Epidemiol.* 1991;8(3):199–208.
8. Sellers TA, Chen PL, Potter JD, Bailey-Wilson JE, Rothschild H, Elston RC. Segregation analysis of smoking-associated malignancies: evidence for Mendelian inheritance. *Am J Med Genet.* 1994;52(3):308–314.
9. Yang P, Schwartz AG, McAllister AE, Swanson GM, Aston CE. Lung cancer risk in families of nonsmoking probands: heterogeneity by age at diagnosis. *Genet Epidemiol.* 1999;17(4):253–273.
10. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics.* 2006;22(1):7–12.
11. Schaid DJ, Batzler AJ, Jenkins GD, Hildebrandt MA. Exact tests of Hardy-Weinberg equilibrium and homogeneity of disequilibrium across strata. *Am J Hum Genet.* 2006;79(6):1071–1080.
12. Ihaka R, Gentleman RR. a language for data analysis and graphics. *J Comput Graph Statist.* 1996;5(3):299–314.
13. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet.* 1955;19(4):251–253.
14. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet.* 2006;78(4):629–644.
15. Lin S, Chakravarti A, Cutler DJ. Exhaustive allelic transmission disequilibrium tests as a new approach to genome-wide association studies. *Nat Genet.* 2004;36(11):1181–1188.
16. Saccone SF, Hinrichs AL, Saccone NL, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet.* 2007;16(1):36–49.
17. Egleton RD, Brown KC, Dasgupta P. Nicotinic acetylcholine receptors in cancer: multiple roles in proliferation and inhibition of apoptosis. *Trends Pharmacol Sci.* 2008;29(3):151–158.
18. Lam DC, Girard L, Ramirez R, et al. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. *Cancer Res.* 2007;67(10):4638–4647.
19. Maneckjee R, Minna JD. Opioid and nicotine receptors affect growth regulation of human lung cancer cell lines. *Proc Natl Acad Sci USA.* 1990;87(9):3294–3298.
20. Maneckjee R, Minna JD. Opioids induce while nicotine suppresses apoptosis in human lung cancer cells. *Cell Growth Differ.* 1994;5(10):1033–1040.
21. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science.* 2002;296(5576):2225–2229.
22. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;21(2):263–265.

Funding

This work was supported in part by the following: National Institutes of Health (NIH) grants U01CA76293 (Genetic Epidemiology of Lung Cancer Consortium), R01CA058554, R01CA093643, R01CA099147, R01CA099187, R01ES012063, R01 ES013340, R03CA77118, R01CA80127, P30ES 06096, P50CA70907 (Specialized Program of Research Excellence), N01HG65404, N01-PC35145, P30CA22453, R01CA63700, DE-FGB-95ER62060, R01CA55769, R01CA121197, R01CA133996, Mayo Clinic intramural research funds, and Department of Defense VITAL grant. This study was supported in part by NIH, the Intramural Research Programs of the National Cancer Institute, and the National Human Genome Research Institute. The authors had full responsibility of design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript.

Notes

Marshall W. Anderson and Ming You contributed equally to this work.

We thank the Fernald Medical Monitoring Program for sharing their biospecimens and data with us. We are grateful to the lung cancer families who participated in this research, and for the high caliber service of Vanderbilt University Microarray Shared Resource, Human Genetics Division Genotyping Core, and to Qiong Chen, Shaw Levy, Victor Guallar, and Jennifer Baker for their assistance in various aspects of this work.

Manuscript received April 2, 2008; revised June 30, 2008; accepted July 2, 2008.

Oncogene (2008), 1–10
© 2008 Nature Publishing Group All rights reserved 0950-9232/08 $30.00
www.nature.com/onc

ONCOGENOMICS

DNA amplification is a ubiquitous mechanism of oncogene activation in lung and other cancers

WW Lockwood[1], R Chari[1], BP Coe[1], L Girard[2], C MacAulay[3], S Lam[3], AF Gazdar[2], JD Minna[2] and WL Lam[1]

[1]*Departments of Cancer Genetics and Developmental Biology, British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada;* [2]*Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, TX, USA and* [3]*Department of Cancer Imaging, British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada*

Chromosomal translocation is the best-characterized genetic mechanism for oncogene activation. However, there are documented examples of activation by alternate mechanisms, for example gene dosage increase, though its prevalence is unclear. Here, we answered the fundamental question of the contribution of DNA amplification as a molecular mechanism driving oncogenesis. Comparing 104 cancer lines representing diverse tissue origins identified genes residing in amplification 'hotspots' and discovered an unexpected frequency of genes activated by this mechanism. The 3431 amplicons identified represent ~10 per hematological and ~36 per epithelial cancer genome. Many recurrently amplified oncogenes were previously known to be activated only by disease-specific translocations. The 135 hotspots identified contain 538 unique genes and are enriched for proliferation, apoptosis and linage-dependency genes, reflecting functions advantageous to tumor growth. Integrating gene dosage with expression data validated the downstream impact of the novel amplification events in both cell lines and clinical samples. For example, multiple downstream components of the EGFR-family-signaling pathway, including *CDK5*, *AKT1* and *SHC1*, are overexpressed as a direct result of gene amplification in lung cancer. Our findings suggest that amplification is far more common a mechanism of oncogene activation than previously believed and that specific regions of the genome are hotspots of amplification.
Oncogene advance online publication, 7 April 2008; doi:10.1038/onc.2008.98

Keywords: gene amplification; array CGH; gene expression; integrative analysis; lung cancer; EGFR signaling

Introduction

Genetic aberration and the consequential activation of oncogenes are key to cancer development. Chromosomal translocation is known as the major event in oncogene activation (Futreal *et al.*, 2004). However, the prevalence of alternate mechanisms, such as DNA amplification, have not been extensively quantified, even though oncogenes have been found in: (i) cytogenetically visible double minutes, which are circular, extrachromosomal elements, a few megabases in size that replicate autonomously, (ii) homogenous staining regions, which are large regions of tandem repeats within a chromosome, thought to be formed by repeated breakage-fusion-bridge cycles and (iii) discrete insertions distributed throughout the genome (Albertson, 2006).

Surprisingly, relatively few oncogenes, when compared to chromosome translocation, have been shown to undergo amplification as mechanism of activation during cancer development. In fact, a recent version (January 22, 2007) of a census of genes causally implicated in cancer (cancer genes) originally described by Futreal *et al.* (2004) reported only seven oncogenes meeting their criteria as being recurrently amplified in the development of human cancers: *AKT2* in ovarian cancer, *ERBB2* in breast and ovarian cancer, *MYCL1* in small cell lung cancer, *MYCN* in neuroblastoma, *REL* in Hodgkin lymphoma, epidermal growth factor receptor (*EGFR*) in glioma and non-small cell lung cancer (NSCLC), and *MYC* in numerous cancers. We propose that the low incidence of oncogenes activated by amplification may be attributed to the failure of detection, rather than governed by tumor biology. Unlike copy number gains, which are generated by aneueploidy or unbalanced translocations and affect large chromosomal regions, amplifications are traditionally defined as the increase of chromosome segments 0.5–10 megabases (Mb) in size (Myllykangas *et al.*, 2006). The small size of amplicons may escape detection by conventional cytogenetic methods; consequently, the contribution of DNA amplification to the oncogenic process may be grossly underestimated. With advances in high-resolution whole-genome-profiling technologies (Tonon *et al.*, 2005; Garnis *et al.*, 2006), the complexity of the cancer genome is becoming evident, and the prevalence of DNA amplification as a mechanism in the activation of oncogenes needs to be re-evaluated.

In this study, we determined the precise boundaries of amplified chromosomal segments in 104 cancer cell lines

---

Correspondence: WW Lockwood, British Columbia Cancer Research Centre, 675 West 10th Aveue, Vancouver, British Columbia, Canada V5Z 1L3.
E-mail: wlockwood@bccrc.ca
Received 4 December 2007; revised 28 January 2008; accepted 1 March 2008

from multiple tissues of origin and deduced novel regions of the genome, which are hotspots for genomic amplification. These hotspots were then analysed for their association with genes involved in tumorigenesis and fragile sites. We assessed the functional impact of a subset of the identified hotspots in a panel of NSCLC cell lines and tumors to determine their effect on gene transcription levels and their contribution to the activation of cellular pathways potentially involved in lung tumorigenesis.

Results

Identification of discrete amplicons in cancer genomes

Twenty-four thousand eight hundred and ninety-two genomic loci were assessed for each of the 104 cell lines, scanning all autosomes at a resolution of ~50 kb (Coe *et al*., 2007). Altogether, 3431 amplicons were detected across all samples (see Supplementary Methods) with an average size of 0.68 Mb and a median of 0.33 Mb (Table 1, Supplementary Table 1). The number of amplicons per genome varied from 0 to 199 with an average of 33. Hematological malignancies (leukemia and lymphomas) had ~10 amplicons, whereas epithelial cancers had an average of 36.

Unexpected frequent amplification of known oncogenes

The most recent version (January 22, 2007) of the Cancer Gene Consensus of the Cancer Genome Project at the Sanger Institute (Futreal *et al*., 2004) contains 363 cancer genes whose aberration are causal in the development of specific cancers. Of these, 70 are tumor suppressor genes, 292 are oncogenes and one can act as both. Only seven (2%) of these oncogenes were shown to be predominately activated by amplification compared to 268 (92%), which are activated mainly by chromosomal translocation. Our data showed amplification at these loci: *MYC* (28/104), *ERBB2* (10/104), *EGFR* (7/104), *MYCL1* (6/104), *AKT2* (3/104) and *MYCN* (1/104). *REL* amplification was not detected in our dataset, as Hodgkin Lymphoma, in which this gene is amplified, is not represented in our study.

Unexpectedly, 145 of the 292 oncogenes (~50%) showed amplification, with 78 oncogenes (27%) at ⩾2 times (Supplementary Table 2). Of the genes amplified in ⩾5 cell lines, only *MYC*, *ERBB2*, *EGFR* and *MYCL1* have been reported. The frequent amplification of *SS18L1*, *NTRK1* and *PRDM16* are novel findings, as translocation was the known mechanism. Indeed, numerous oncogenes, which are primarily activated by translocation, were commonly amplified in the sample set (Supplementary Table 3). The number of oncogenes amplified per genome also varied with an average of 3.5 genes. Remarkably, the genomes of NSCLC HCC1195 (Supplementary Figure 1) and SCLC line H526, each harbor 22 amplified oncogenes, whereas 25 of the lines had no known oncogenes amplified (Supplementary Table 4).

Novel hotspots of frequent genomic amplification in cancer genomes

The high incidence of oncogene amplification per genome suggested that this is a common mechanism of gene activation. Therefore, the discovery of genomic regions that undergo frequent copy number amplification may lead to the identification of novel oncogenes. The genomic coordinates of all amplicons were determined and aligned for all 104 samples (Figure 1). DNA segments amplified ⩾5 times were stringently considered as hotspots; they are found in ~5% of samples (see Supplementary Methods). In total, 135 hotspots covering 3% of genome were identified with an average size of 0.67 Mb. Regions of genomic amplification were distributed on all autosomes except chromosome 4, and in all tumor types analysed (Supplementary Table 5). Amplicons are most frequently localized to 1q21–23, 5p15, 7p13–11, 8q22–24, 11q13, 14q12–21, 14q32, 17q12–21 and 20q13. A total of 538 unique genes were contained within the hotspots (Supplementary Table 6) (see Supplementary Methods). Interestingly, the majority of these hotspots did not contain the 292 known oncogenes.

Table 1 Summary and distribution of amplicons by cancer type

Tissue type	*Number of Samples*	*Total number of amplicons*	*Amplicons/ Tumor*	*Average size (Mb)*	*Most frequent amplification*[a]
Lung	53	1690	31.9	0.68	8q24.21 (28%)
Breast	17	905	53.2	0.77	8q24.21 (59%)
Lymphoid	9	101	11.2	0.54	9p13.3, 13q31.3, 18q21.33–22.1 (33%)
Cervix	8	45	5.6	0.65	5p15.33 (38%)
Skin	4	395	98.8	0.35	7p13.3 & 7q35 (75%)
Blood	3	23	7.7	0.80	N/A
Prostate	3	82	27.3	0.86	14q21.3–22.1 (67%)
Bone	2	104	52	1.02	N/A
Colon	2	22	11	0.70	N/A
Ovary	2	3	1.5	0.33	N/A
Liver	1	61	61	0.83	N/A
Total	104	3431	32.8	0.68	

[a] The most frequent regions were determined only if found in >2 samples for the corresponding tissue type.

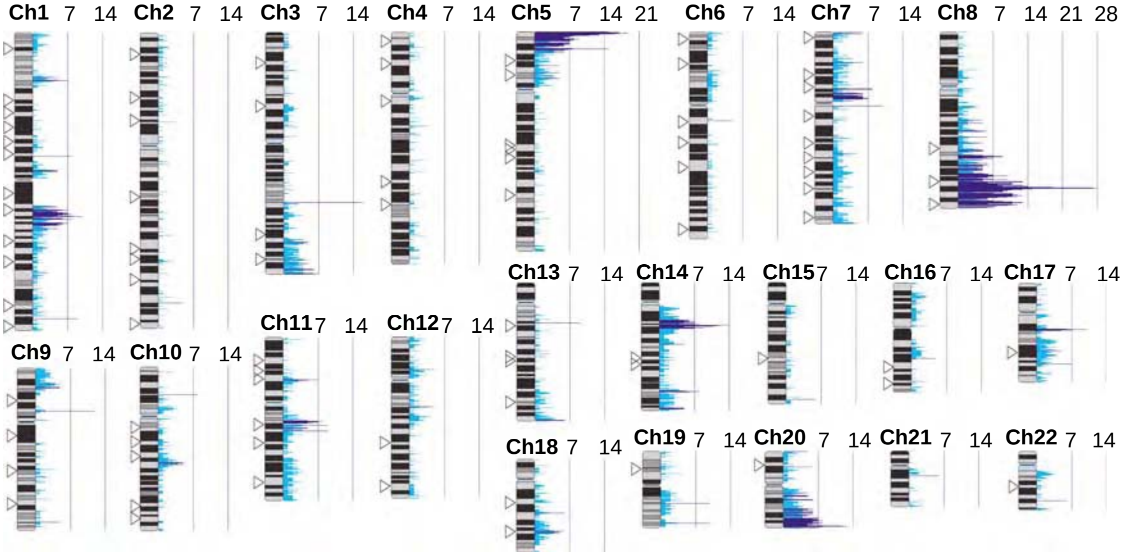


Figure 1 Hotspots of amplification in cancer genomes. A histogram summarizing the regions of amplification across all 104 samples with the resulting values scaled to the segment with the highest count (28) and plotted against their corresponding genomic position. Hotspots are denoted by the dark blue shading, whereas the light blue shading represents regions amplified $\leqslant 5$ times. Triangles mark common fragile sites. Detailed genomic position of hotspots and common fragile sites are provided in Supplementary Tables 5 and 17.

There was no association between amplification hotspots and known fragile sites in the human genome based on χ^2 test at the chromosome band level, even though colocalization does exist, for example, the three fragile sites on 8q. Figure 1 summaries the location of the 86 common fragile sites assayed relative to hotspots of amplification.

Novel amplification hotspots contain putative oncogenes

Remarkably, 27 of the top 100 most frequently amplified genes have been previously described to be overexpressed in various cancers (Supplementary Table 7), but aside from *MYC* and *ERRB2*, the mechanism leading to overexpression of these genes was largely unknown. To further explore the properties of the amplified genes, functional and biological characteristics were evaluated through the use of Ingenuity Pathways Analysis (Ingenuity Systems, see (see Supplementary Methods). Functional analysis identified a significant association between the amplified genes with genes involved in cancer ($P = 6.67\text{E}{-}06$–$3.03\text{E}{-}02$; the two significance values refer to a range of specific sub-functions) and other diseases (Supplementary Tables 8 and 9). Furthermore, canonical pathway analysis was used to determine the main signaling pathways in which the amplified genes were involved (Supplementary Table 10). Neuregulin Signaling ($P = 1.12\text{E}{-}02$), also known as EGFR-family-signaling, was the most significantly affected with *GRB7*, *SHC1*, *SRC*, *EGFR*, *ERBB2* and *AKT1* comprising the amplified genes.

Impact of amplification on gene expression levels

To understand the effects of amplification on gene regulation and transcription, we focused on one type of cancer. Parallel gene expression profiles and array comparative genomic hybridization (CGH) data were integrated for 27 NSCLC cell lines. The expression levels for genes within amplification hotspots (displayed in Supplementary Figure 2) were compared between samples with amplification and those with neutral copy number status using the Mann–Whitney *U* test (see Supplementary Methods). In total, 221 out of 442 of the amplified genes were expressed at significantly higher levels ($P \leqslant 0.05$) with increased gene dosage (Figure 2 and Supplementary Table 11). For the majority of these genes, amplification is a novel mechanism for activation, although the expression levels of a subset, such as *MYC, EGFR, CDK4*, *MAFB* and *MET*, are known to be affected by increase in gene dosage.

Multiple components of the EGFR-family-signaling pathway are activated by DNA amplification in NSCLC cell lines and clinical tumors

To relate the genes activated by amplification in NSCLC to biological functions, Functional and Canonical Pathway Analysis were performed using IPA software (Supplementary Tables 12 and 13). EGFR-family-signaling was the most affected canonical pathway ($P = 6.03\text{E-}03$) with five genes: *AKT1*, *CDK5*, *EGFR*, *MYC* and *SHC1* amplified and overexpressed in the 27 NSCLCs (Table 2, Figure 2). The amplification and subsequent overexpression of *AKT1*, *CDK5* and *SHC1* are novel findings in NSCLC. Figure 3 displays the interaction of these genes during EGFR-family-signaling and the resulting downstream effects of the activation of this pathway, which includes cell proliferation and survival. Interestingly, nearly 60% of the cell lines analysed had one or more components of the EGFR family pathway overexpressed as a result of amplification. The alteration of *EGFR* and *MYC* alone

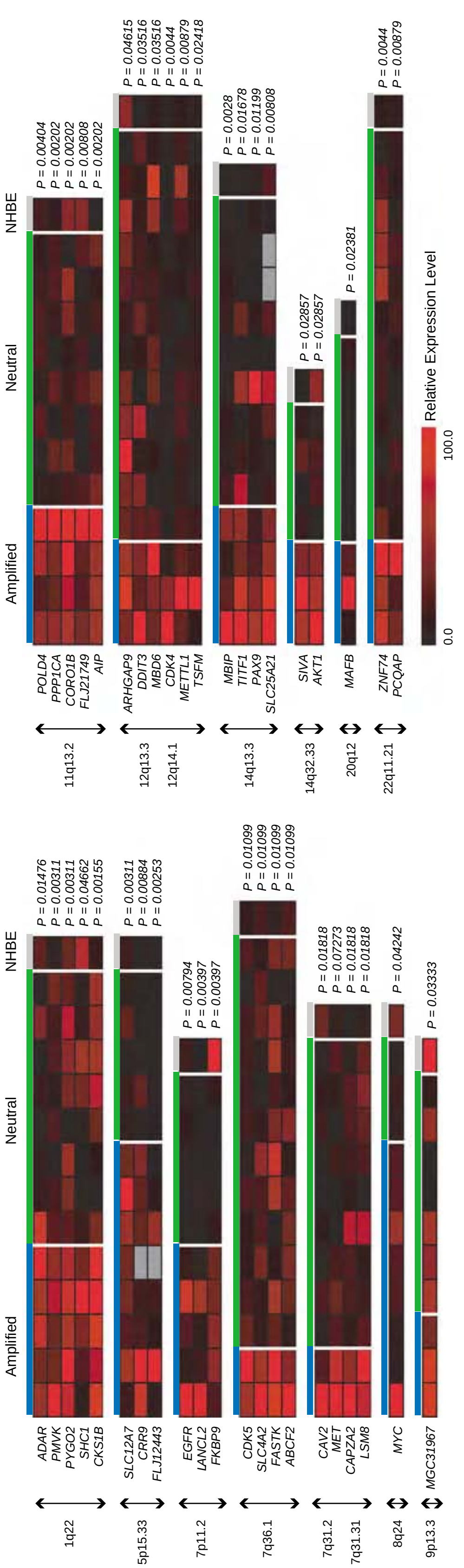


Figure 2 Impact of amplification on gene transcription levels. The relative expression values for samples with amplification and those with neutral copy number status are plotted as heatmaps for overexpressed genes from representative hotspots. The expression values for each gene have been normalized and scaled across the samples from 0 to 100.

Table 2 Canonical pathways affected by amplification in non-small cell lung cancer

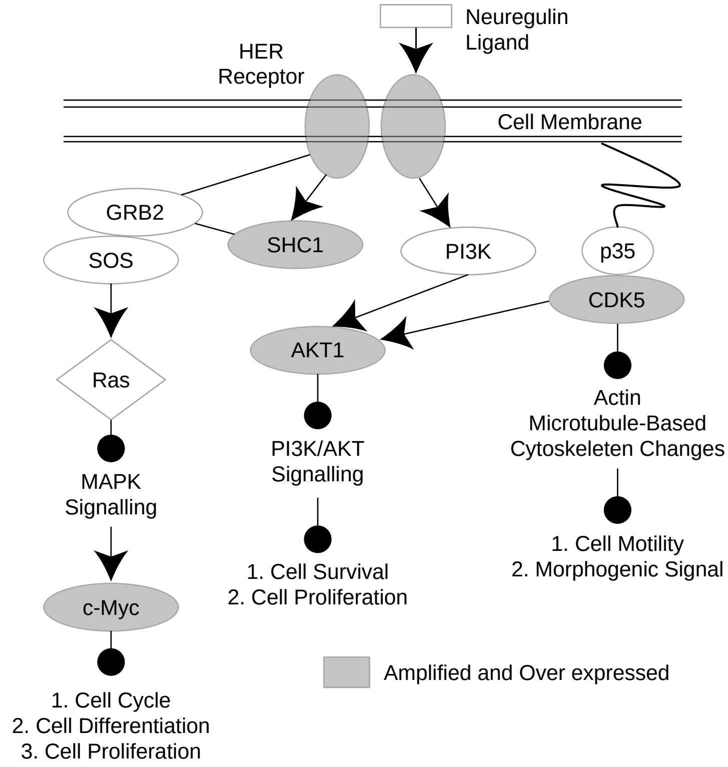
Signaling pathway	P-*value*	*Genes*
Neuregulin	6.0E−03	*MYC, CDK5, SHC1, EGFR* and *AKT1*
Huntington's disease	1.1E−02	*SDHA, POLR2H, CDK5, SHC1, POLR2J, EGFR* and *AKT1*
VEGF	3.1E−02	*VEGF, ARNT, SHC1* and *AKT1*
Insulin receptor	3.2E−02	*PARD3, PPP1R3D, PPP1CA, SHC1* and *AKT1*
Nitric oxide	3.9E−02	*VEGF, CALM1* and *AKT1*
Hypoxia	4.3E−02	*VEGF, ARNT* and *AKT1*

could not explain pathway disruption in all cases as ~31% (5/16) of samples with activated downstream components harbored amplification of one or more of *CDK5*, *SHC1* or *AKT1* independent of *EGFR* and *MYC*. Furthermore, by breaking the NSCLC lines down into their histological subtypes, it was discovered that 14 out of 20 (70%) adenocarcinoma samples—whereas no squamous and only one large cell carcinoma samples—had altered components, suggesting that the disruption of this pathway is prevalent in the adenocarcinoma subtype of lung cancer.

To further validate our results, quantitative real-time PCR was performed on select genes, *AKT1*, *CDK5* and *SHC1*. First, expression levels for these genes were determined using the $\Delta\Delta C_t$ method and compared between cell lines and normal lung tissue to confirm their overexpression. Relative to the normal lung reference, *AKT1* was 10.49-fold overexpressed in samples with gene amplification compared to 1.94-fold overexpression in NSCLC cells with neutral copy number status for this gene (Supplementary Figure 3). Likewise, *CDK5* (Supplementary Figure 4) and SHC1 (Figure 4) also showed higher expression with increase gene dosage, suggesting a strong correlation of gene dosage and transcription levels for these genes. Expression changes held true in clinical specimens as clinical adenocarcinoma samples frequently showed overexpression of *AKT1, CDK5* and *SHC1* compared to their corresponding matched normal lung tissues (Supplementary Figure 3c, 4c and Figure 4c, respectively). Since these genes were hypothesized to be overexpressed due to DNA amplification, a one-tailed Wilcoxon sign-rank test was used to determine whether overexpression of these genes was significant in the set of matched tumor and normal samples. Indeed, each gene was significantly overexpressed in the tumors compared to their matched normal ($P<0.05$), confirming the results from the cell lines (qPCR data is provide in Supplementary Tables 14 and 15).

Discussion

Oncogene activation is traditionally associated with translocation events. We hypothesized that DNA amplification is a prevalent, but underestimated,


Neuregulin Ligand
HER Receptor
Cell Membrane
GRB2
SOS
SHC1
PI3K
p35
CDK5
AKT1
Ras
MAPK Signalling
PI3K/AKT Signalling
Actin Microtubule-Based Cytoskeleten Changes
1. Cell Motility
2. Morphogenic Signal
1. Cell Survival
2. Cell Proliferation
c-Myc
Amplified and Over expressed
1. Cell Cycle
2. Cell Differentiation
3. Cell Proliferation

Gene	# of Samples with Amplification and Overexpression		% Samples (n=27)
EGFR	5	H1819, HCC2279, H3255, HCC4006, HCC827	18.5
SHC1	5	H1395, H1993, H2122, HCC1195, HCC366	18.5
CDK5	2	H2009, HCC193	7.4
AKT1	3	HCC366, H1819, HCC461	11.1
c- MYC	8	H1395, H2122, HCC1195, HCC2279, HCC827, H2087, H1975, H460	29.6
*Total	16		59.3

Figure 3 Frequent amplification and overexpression of multiple EGFR-family-signaling components in non-small cell lung cancer (NSCLC). Diagram highlighting the interaction of EGFR, SHC1, CDK5, SHC1 and MYC in the EGFR family-signaling pathway. Altered components are shaded grey. The table summarizes the number and specific samples with amplification and overexpression of each pathway component. The total represents the number of samples with at least one pathway component amplified and overexpressed.

mechanism of oncogene activation in cancer genomes. To our knowledge, no studies to date have assembled a large panel of paired high-resolution copy number and gene expression data to accurately assess this question. In this study, we examined 104 cancer cell lines comprising various tissues of origin (Table 1) at 24 892 autosomal loci per genome, a resolution that detected amplicons as small as 0.05 Mb in size (Ishkanian *et al*., 2004) and discovered that not only is the incidence of oncogene amplification much greater than previously believed, but specific regions of the genome are hotspots for segmental amplification in cancer cells.

Amplification as a major mechanism of oncogene activation

The activation of oncogenes is a hallmark of tumor development. Cancer cells frequently display chromosome rearrangements resulting in the deregulation of gene expression, as well as in the fusion of genes raising oncogenic activity. As such, the majority of known oncogenes, including 92% of those analysed in this study, have been discovered through their involvement in disease-specific chromosomal translocations (Futreal *et al*., 2004). Thus, the high incidence of amplification we report suggests that oncogenes may have multiple mechanisms of activation, with the increase in gene dosage being a prominent mechanism of activation. This was particularly evident in the fact that genes, which have been shown to be activated primarily by translocation, were frequently amplified (Supplementary Table 3). The majority of these genes has not been shown to be activated by amplification previously, and as such, this data represent a novel finding. For example, t(14;20)(q32;q12) translocation is known to juxtapose *IgH* enhancers to the *MAFB* gene locus upregulating its expression in multiple myeloma (Wang *et al*., 1999; Boersma-Vreugdenhil *et al*., 2004). We

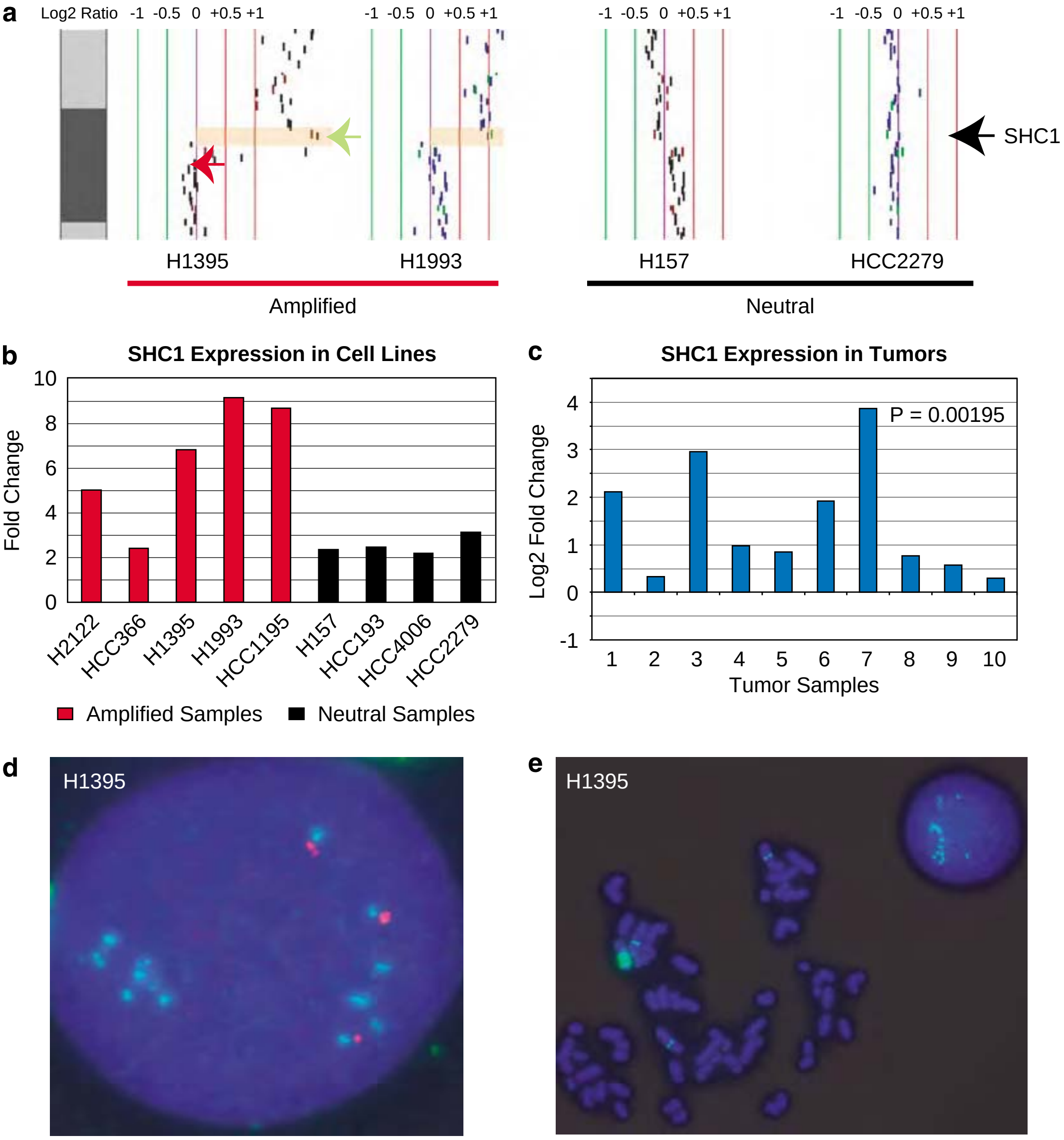


Figure 4 SHC1disruption in NSCLC cell lines and clinical tumors. (**a**) Representative array CGH profiles for samples with and without *SHC1* amplification. Vertical lines denote $\log_2$ signal ratios from −1 to 1 with copy number increases to the right (red lines) and decreases to the left (green lines) of 0 (purple line), with amplified region shaded orange. Red and green arrows mark clones used in subsequent Fluorescence *in situ* hybridization (FISH) analysis. (**b**) *SHC1* expression in NSCLC cell lines. The normalized fold change of expression compared to a normal lung reference is plotted for samples with amplification (red) and those with neutral copy number status (black). (**c**) Overexpression of *SHC1* in clinical tumors. The $\log_2$ fold change in expression levels of *SHC1* relative to their matched normal lung tissue is plotted for each tumor. The *P*-value from the Wilcoxon sign-rank test is indicated. (**d** and **e**) FISH confirmation of *SHC1* amplification in H1395. FISH was performed using BAC clones mapping to *SHC1* (RP11-624P9) and to an adjacent neutral copy number region (RP11-313J15).

demonstrated that *MAFB* amplification also occurs in lung (H1395, H1650 and H1666), cervical (SW756) and liver (HepG2) cancer cells. Likewise, *NTRK1* is an oncogene frequently activated by translocation (Roccato *et al*., 2005). Fusion with *TPR*, *TPM3* or *TFG* results in constitutive tyrosine kinase activity (Pierotti *et al*., 1996). Remarkably, we also detected *NTRK1* copy number increased in lung (HCC366, HCC1833, HCC1195, H82, H526, H2122 and H187) and breast (ZR7530) cancer genomes, suggesting that amplification and subsequent overexpression may be an alternate mechanism of activation.

Existence of an amplifier phenotype

There is evidence suggesting that some cancer cells have a greater propensity to undergo DNA amplification than others and that there is an underlying genetic basis for this 'amplifier phenotype' (Albertson, 2006). Our data showed that the number of oncogenes amplified may differ in individual genomes. The variation existed across both general cancer classes and individual tissue types. In addition, amplification of the same genes in the different tumor types suggests that there may be a selective advantage to have certain genes or their related functions elevated in the context of cancer development and that amplifications are not simply byproducts of general genomic instability characteristic of late-stage tumors. It was also common for a sample to simultaneously harbor multiple amplicons on different chromosomes, highlighting the possibility of an underlying genetic basis for amplification development (Supplementary Figure 1). It has been proposed that amplifications are mainly related to solid tumors and are seldom involved in hematological malignancies in which oncogene activation is generally associated with translocations (Mitelman, 2000). Indeed, only 11 known oncogenes were amplified in leukemia or lymphoma genomes in our data set, whereas 144 were amplified in epithelial cases (Supplementary Table 2). Similar results were observed in the number of amplicons in each genome ($\sim$10 per hematological, $\sim$36 in epithelial samples). These results suggest that subsets of cancers, such as the epithelial cancers, are driven by an amplifier phenotype, whereas others, typically hematological malignancies, develop mainly through different genetic mechanisms, such as chromosomal translocation.

Novel amplification hotspots are enriched for putative oncogenes

Figure 1 indicates that regions of frequent copy number amplification are preferentially localized in genome. These results are similar to those found in a bibliomics survey, which looked across 73 different neoplasms using conventional CGH (Myllykangas *et al*., 2006). However, in this study, we further refined the regions beyond the chromosome band level, determining the exact genes affected by these aberrations (Supplementary Table 6), identifying novel hotspots, such as the discrete amplicons on 14q.

There are two potential factors that may determine the localization of amplification hotspots. First, the selective pressure imposed on the tumor may lead to the selection of amplification of regions containing genes advantageous to tumor growth. Consistent with this theory, we observed a significant enrichment for genes involved in cellular functions and canonical pathways commonly involved in tumorigenesis within the amplification hotspots (Supplementary Tables 8–10). Many genes implicated in key biological processes, such as cell cycle and cell growth and proliferation contained within these regions that may be considered as novel candidate oncogenes. In addition, 27 of the top 100 genes within the hotspots have been previously described to be overexpressed in cancer further supporting their oncogenic role (Supplementary Table 7). Since the mechanism leading to the overexpression of the majority of these genes was previously unknown, our data suggest that amplification may be a key mechanism of their activation.

Second, the intrinsic features of the chromosome regions themselves may be involved in their preferential amplification (Wahl *et al*., 1984). Mechanistic models, such as breakage-fusion-bridge and episome excision, imply that two double-stranded DNA breaks are required to initiate amplification generation (Myllykangas and Knuutila, 2006). As such, it has been proposed that regions of the genome that are more susceptible to breakage have a greater propensity to undergo amplification, such as, fragile sites (Hellman *et al*., 2002; Buttel *et al*., 2004). Conventional CGH studies indicated that many amplification hotspots colocalized with fragile sites, however, this association was not statistically significant on a genome-wide scale, presumably due to inadequate resolution (with chromosome band level comparison) as fragile sites and amplification hotspots covered 30 and 45% of the genome, respectively (Myllykangas *et al*., 2006). Our analysis addressed this problem as the increased resolution of the platform used in this study allowed the refinement of amplification hotspots, limiting their coverage to $\sim$3% of the genome. However, although we observed a general trend of the colocalization of amplification hotspots and common fragile sites (Figure 1), the global association was still not significant. We speculate that the cloning of fragile sites to determine their specific sequences is needed to complement our array CGH data to accurately assess their association. Nevertheless, there is a strong possibility that the hotspot regions represent damage-prone sites in the genome and further investigation in the future is warranted. Notably, in addition to fragile sites, other genomic features, such as copy number variations and segmental duplications, may also contribute to DNA rearrangement in cancer cells (Squire *et al*., 2003).

Global impact of amplification on gene expression levels in NSCLC

Although array CGH allows the fine mapping of amplification boundaries at unprecedented resolution, multiple genes may map to an individual amplicon. Therefore, integration of copy number and expression data is needed to distinguish overexpressed genes from bystander genes within amplicons. To accomplish this, we integrated parallel array CGH and expression data for a subset of 27 NSCLC cell lines. Aside from genes known to be activated by amplification in NSCLC, such as *EGFR* and *MYC*, our data suggest that the expression of oncogenes, including *CDK4*, *MAFB*, *BCL11B* and *MET*, is also driven by amplification (Figure 2). Since these genes are typically activated by translocation or missense mutation in malignancies other than NSCLC, these data further supports our hypothesis that amplification is an alternate mechanism of oncogene activation in a subset of cancers.

The integration of the data sets identified expressed genes within novel amplification hotspots in lung cancer that are potentially involved in tumorigenesis. For example, thyroid transcription factor 1 (*TITF1*) in the novel hotspot at 14q12–q13 is known to be overexpressed specifically in lung adenocarcinoma, the predominate subtype of NSCLC from which the 27 cell lines in our study were derived (Fabbro *et al*., 1996). This gene encodes a homeodomain transcription factor that is involved in regulating pulmonary development and gene expression (Apergis *et al*., 1998) and has been proposed to be a lineage marker for tumors arising from the peripheral airway (Stenhouse *et al*., 2004). Adenocarcinomas that express TITF1 are dependent on its persistent expression for survival (Tanaka *et al*., 2007). Acquired somatic alteration of this gene during differentiation leads to aberrant lineage-survival pathway signaling. The resulting tumors become addicted to persistent expression of the gene for survival, known as 'lineage addiction' (Garraway and Sellers, 2006). Our data suggests that amplification may be a mechanism driving the expression of linage-survival oncogenes in a subset of cancers and further supports a role for *TITF1* in lung adenocarcinoma tumorigenesis (Figure 2 and Supplementary Figure 5).

On a global scale, we found amplification had a strong impact on transcription levels as 50% of the genes from amplification hotspots within these samples showed enhanced expression as a consequence of alteration. This falls within observations of previous studies that reported 19.3–62% of amplified genes being overexpressed (Hyman *et al*., 2002; Pollack *et al*., 2002; Wolf *et al*., 2004; Heidenblad *et al*., 2005). Nevertheless, since this is the first study to integrate high-resolution copy number and gene expression profiles in lung cancer on a whole-genome scale, future analysis will be needed to confirm the functional impact of the genes in each amplicon.

Novel disruptions of the EGFR-family-signaling pathway in NSCLC by gene amplification

Our data indicate that multiple components of the EGFR-family-signaling pathway are frequently amplified and overexpressed in NSCLC (Figure 3). Deregulation of this network commonly occurs in cancer and specifically NSCLC. In NSCLC, the mechanism of deregulation is usually attributed to receptor overexpression or point mutations in the catalytic domain, resulting in ligand-independent constitutive receptor activation and signaling (Bublil and Yarden, 2007). However, in cases with normal levels of wild-type receptor, aberrant constitutive signaling may also occur. Strikingly, our data suggest that amplification of downstream signaling components may be an alternate mechanism of pathway activation in a subset of tumors. Although alteration of this pathway was detected in 59% of the cell lines analysed, only 31% of these cases could be explained by *EGFR* activation. Indeed, the majority of cell lines with pathway perturbation (69%) contained amplification of key signaling components downstream of the receptor level (Figure 3). The overexpression of these components was confirmed in NSCLC tumors, highlighting their clinical significance (Figure 4 and Supplementary Figures 3 and 4). Although *MYC* amplification has been previously reported, this is the first study to describe the frequent amplification and overexpression of *SHC1*, *AKT1* and *CDK5* in NSCLC. Since these genes are involved in the activation of mitogenic-signaling pathways involved in EGFR-induced transformation and their overexpression has been previously implicated in cancer, these results suggest that their direct genetic activation may play a causal role in NSCLC tumorigenesis and highlights the impact of these novel amplifications on cancer biology (Hennessy *et al*., 2005).

The direct amplification of downstream components may also have substantial effect on the response to clinical treatment strategies. The high frequency of *EGFR* family overexpression in cancer has led to the development of targeted therapeutics aimed at inhibiting receptor function. For example, anti-ErbB2 antibodies are currently used for breast cancer treatment and EGFR-specific tyrosine kinase inhibitors, such as Gefitinib and Erlotinib are used in NSCLC therapy (Bublil and Yarden, 2007). The receptor-independent activation of downstream signaling components would impact the effectiveness of these treatment strategies as constitutive activation of signaling pathways would occur regardless of receptor inhibition. Previous studies have shown that downregulation of the AKT/PI3K signaling pathway is required for EGFR tyrosine kinase inhibitors to induce apoptosis in cancer cells (Hemstrom *et al*., 2006). In addition, activated AKT/PI3K signaling due to *MET* amplification has been shown to lead to Gefitinib resistance in NSCLC cells (Engelman *et al*., 2007) Thus, the direct activation of AKT1 and CDK5 would lead to resistance in a NSCLC due to maintained AKT/PI3K signaling in the presence of inhibitor. Likewise, in a drug-resistant NSCLC cell line, alterations of adaptor-protein-mediated signal transduction from EGFR, such as those initiated by SHC1, has been proposed as a possible mechanism of resistance to Gefitinib (Koizumi *et al*., 2005). Our findings highlight the need to assess the activation status of downstream signaling components and suggest that amplification *AKT1*, *SHC1*, *CDK5* and *MYC* may be used for this process.

Conclusion

Since alterations at the DNA level potentially represent causal events in the development of cancer, the genes deregulated as a result of amplification in NSCLC can be viewed as the primary oncogenic targets in a tumor that lead to downstream pathway abrogation. The gain-of-function effect of gene amplifications makes them ideal targets for therapeutic intervention due to the direct nature of their activation and the fact that a tumor can become addicted to their enhanced expression (Weinstein, 2002). Although cancer cell lines were used in this study, the genomic and transcriptional characteristics of such models have been shown to mirror primary tumors and are appropriate systems to identify molecular features that predict or indicate response to targeted therapies (Neve *et al*., 2006; Greshock *et al*., 2007). Furthermore, validation of a subset of amplified genes in primary tumors confirmed their clinical importance. Future studies of additional primary tumors will be required to further validate the role of the hotspots in clinical specimens and confirm that they are not artifacts of *in vitro* culture. Our discovery of high incidence of amplification suggests that it is a major mechanism of oncogene activation in cancer and will provide essential starting points for the discovery of novel oncogenes.

Materials and methods

Whole-genome profiling

DNA copy number profiles for 104 cancer cell lines of lung, breast, prostate, cervical, skin, ovarian, liver and hematological origins were used in this study (Supplementary Table 16). DNA was isolated by proteinase K digestion followed by phenol–chloroform extraction. Array hybridization was performed as previously described (Lockwood *et al*., 2007), using SMRT array v.2 (Ishkanian *et al*., 2004; Watson *et al*., 2007). Array images were analysed using SoftWoRx Tracker Spot Analysis software (Applied Precision, Issaquah, WA, USA). Systematic biases were removed using the stepwise normalization procedure CGH Norm (Khojasteh *et al*., 2005). SeeGH software allowed visualization of $\log_2$ ratio plots in karyograms (Chi *et al*., 2004). All raw array data files have been made publicly available through the System for Integrative Genomic Microarray Analysis (SIGMA), which can be accessed at http://sigma.bccrc.ca (Chari *et al*., 2006).

Gene expression profiling

RNA samples from 27 NSCLC cell lines and normal human bronchial epithelial cells were analysed using the Affymetrix Gene Chips HG-U133A and HG-U133B (Henderson *et al*., 2005; Zhou *et al*., 2006) (Supplementary Table 16). These arrays together represent 23 583 unique genes based on Unigene build 173. The identity of these cell lines has been verified by DNA fingerprint, using the Powerplex 1.2 system (Promega). Data normalization and microarray analysis was performed using Affymetrix Microarray Suite 5.0 as described previously(Zhou *et al*., 2006). The microarray data have been uploaded to GEO (Gene Expression Omnibus, accession number GSE-4824).

Statistical analysis of array data

Detailed methods describing the statistical analysis used for the identification of amplicons, amplification hotspots, amplification hotspots and fragile site colocalization, functional assessment of amplified genes and integration of genomic, and gene expression data are provided in Supplementary Methods.

Gene-specific quantitative real-time reverse transcriptase PCR analysis

TaqMan gene expression assays (AKT1 (Hs00178289_m1), SHC1 (Hs00427539_m1), CDK5 (Hs00358991_g1) and 18 s rRNA (Hs99999901_s1) were performed using 100 ng of cDNA samples in a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). The $\Delta\Delta C_t$ method was used for expression quantification using the average cycle threshold of 18S rRNA for normalization (Coe *et al*., 2006) and human lung total RNA (AM7968, Ambion, Austin, TX, USA) as a reference. For clinical samples, total RNA was isolated from 10 microdissected frozen lung adenocarcinoma and matched normal tissue obtained from Vancouver General Hospital using RNeasy Mini Kits (QIAGEN Inc., Mississauga, ON, USA) and 1 μg was converted to cDNA for gene-specific quantitative PCR for *AKT1*, *SHC1*, *CDK5* and *18S* rRNA. Cycle thresholds comparison yielded expression changes in the tumors. Because these genes were hypothesized to be overexpressed, owing to DNA amplification, a one-tailed Wilcoxon sign-rank test was used to determine whether overexpression was significant in the set of matched tumor and normal samples.

Fluorescence in situ *hybridization*

Fluorescence *in situ* hybridization was performed as described previously (Watson *et al*., 2004). Briefly, 100 ng of linker-mediated PCR-amplified BAC DNA was labeled through a random priming reaction with Spectrum Green or Red dUTP (Vysis, Markham, ON, USA). Hybridization was performed in a 50% formamide buffer at 37 °C for 18 h and imaged with Q Capture imaging software (Q Imaging, Burnaby, BC, USA).

Acknowledgements

This work was supported by funds from CIHR, Genome Canada/BC, Lung Cancer SPORE P50CA70907, DOD VITAL, the Gillson Longenbaugh and Anderson Charitable Foundations as well as scholarships from NSERC, CIHR and MSFHR to WWL, RC and BPC.

Data deposition: Gene Expression Omnibus, accession number GSE-4824.

References

Albertson DG. (2006). Gene amplification in cancer. *Trends Genet* **22**: 447–455.

Apergis GA, Crawford N, Ghosh D, Steppan CM, Vorachek WR, Wen P *et al.* (1998). A novel nk-2-related transcription factor associated with human fetal liver and hepatocellular carcinoma. *J Biol Chem* **273**: 2917–2925.

Boersma-Vreugdenhil GR, Kuipers J, Van Stralen E, Peeters T, Michaux L, Hagemeijer A *et al.* (2004). The recurrent translocation t(14;20)(q32;q12) in multiple myeloma results in aberrant expression of MAFB: a molecular and genetic analysis of the chromosomal breakpoint. *Br J Haematol* **126**: 355–363.

Bublil EM, Yarden Y. (2007). The EGF receptor family: spearheading a merger of signaling and therapeutics. *Curr Opin Cell Biol* **19**: 124–134.

Buttel I, Fechter A, Schwab M. (2004). Common fragile sites and cancer: targeted cloning by insertional mutagenesis. *Ann NY Acad Sci* **1028**: 14–27.

Chari R, Lockwood WW, Coe BP, Chu A, Macey D, Thomson A *et al.* (2006). SIGMA: a system for integrative genomic microarray analysis of cancer genomes. *BMC Genomics* **7**: 324.

Chi B, DeLeeuw RJ, Coe BP, MacAulay C, Lam WL. (2004). SeeGH—a software tool for visualization of whole genome array comparative genomic hybridization data. *BMC Bioinformatics* **5**: 13.

Coe BP, Lockwood WW, Girard L, Chari R, Macaulay C, Lam S *et al.* (2006). Differential disruption of cell cycle pathways in small cell and non-small cell lung cancer. *Br J Cancer* **94**: 1927–1935.

Coe BP, Ylstra B, Carvalho B, Meijer GA, Macaulay C, Lam WL. (2007). Resolving the resolution of array CGH. *Genomics* **89**: 647–653.

Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO *et al.* (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**: 1039–1043.

Fabbro D, Di Loreto C, Stamerra O, Beltrami CA, Lonigro R, Damante G. (1996). TTF-1 gene expression in human lung tumours. *Eur J Cancer* **32A**: 512–517.

Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R *et al.* (2004). A census of human cancer genes. *Nat Rev Cancer* **4**: 177–183.

Garnis C, Lockwood WW, Vucic E, Ge Y, Girard L, Minna JD *et al.* (2006). High resolution analysis of non-small cell lung cancer cell lines by whole genome tiling path array CGH. *Int J Cancer* **118**: 1556–1564.

Garraway LA, Sellers WR. (2006). Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* **6**: 593–602.

Greshock J, Nathanson K, Martin AM, Zhang L, Coukos G, Weber BL *et al.* (2007). Cancer cell lines as genetic models of their parent histology: analyses based on array comparative genomic hybridization. *Cancer Res* **67**: 3594–3600.

Heidenblad M, Lindgren D, Veltman JA, Jonson T, Mahlamaki EH, Gorunova L *et al.* (2005). Microarray analyses reveal strong influence of DNA copy number alterations on the transcriptional patterns in pancreatic cancer: implications for the interpretation of genomic amplifications. *Oncogene* **24**: 1794–1801.

Hellman A, Zlotorynski E, Scherer SW, Cheung J, Vincent JB, Smith DI *et al.* (2002). A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* **1**: 89–97.

Hemstrom TH, Sandstrom M, Zhivotovsky B. (2006). Inhibitors of the PI3-kinase/Akt pathway induce mitotic catastrophe in non-small cell lung cancer cells. *Int J Cancer* **119**: 1028–1038.

Henderson LJ, Coe BP, Lee EH, Girard L, Gazdar AF, Minna JD *et al.* (2005). Genomic and gene expression profiling of minute alterations of chromosome arm 1p in small-cell lung carcinoma cells. *Br J Cancer* **92**: 1553–1560.

Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. (2005). Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat Rev Drug Discov* **4**: 988–1004.

Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E *et al.* (2002). Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* **62**: 6240–6245.

Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP *et al.* (2004). A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* **36**: 299–303.

Khojasteh M, Lam WL, Ward RK, MacAulay C. (2005). A stepwise framework for the normalization of array CGH data. *BMC Bioinformatics* **6**: 274.

Koizumi F, Shimoyama T, Taguchi F, Saijo N, Nishio K. (2005). Establishment of a human non-small cell lung cancer cell line resistant to gefitinib. *Int J Cancer* **116**: 36–44.

Lockwood WW, Coe BP, Williams AC, MacAulay C, Lam WL. (2007). Whole genome tiling path array CGH analysis of segmental copy number alterations in cervical cancer cell lines. *Int J Cancer* **120**: 436–443.

Mitelman F. (2000). Recurrent chromosome aberrations in cancer. *Mutat Res* **462**: 247–253.

Myllykangas S, Himberg J, Bohling T, Nagy B, Hollmen J, Knuutila S. (2006). DNA copy number amplification profiling of human neoplasms. *Oncogene* **25**: 7324–7332.

Myllykangas S, Knuutila S. (2006). Manifestation, mechanisms and mysteries of gene amplifications. *Cancer Lett* **232**: 79–89.

Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T *et al.* (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**: 515–527.

Pierotti MA, Bongarzone I, Borello MG, Greco A, Pilotti S, Sozzi G. (1996). Cytogenetics and molecular genetics of carcinomas arising from thyroid epithelial follicular cells. *Genes Chromosomes Cancer* **16**: 1–14.

Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE *et al.* (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99**: 12963–12968.

Roccato E, Bressan P, Sabatella G, Rumio C, Vizzotto L, Pierotti MA *et al.* (2005). Proximity of TPR and NTRK1 rearranging loci in human thyrocytes. *Cancer Res* **65**: 2572–2576.

Squire JA, Pei J, Marrano P, Beheshti B, Bayani J, Lim G *et al.* (2003). High-resolution mapping of amplifications and deletions in pediatric osteosarcoma by use of CGH analysis of cDNA microarrays. *Genes Chromosomes Cancer* **38**: 215–225.

Stenhouse G, Fyfe N, King G, Chapman A, Kerr KM. (2004). Thyroid transcription factor 1 in pulmonary adenocarcinoma. *J Clin Pathol* **57**: 383–387.

Tanaka H, Yanagisawa K, Shinjo K, Taguchi A, Maeno K, Tomida S *et al.* (2007). Lineage-specific dependency of lung adenocarcinomas on the lung development regulator TTF-1. *Cancer Res* **67**: 6007–6011.

Tonon G, Wong KK, Maulik G, Brennan C, Feng B, Zhang Y *et al.* (2005). High-resolution genomic profiles of human lung cancer. *Proc Natl Acad Sci USA* **102**: 9625–9630.

Wahl GM, Robert de Saint Vincent B, DeRose ML. (1984). Effect of chromosomal position on amplification of transfected genes in animal cells. *Nature* **307**: 516–520.

Wang PW, Eisenbart JD, Cordes SP, Barsh GS, Stoffel M, Le Beau MM. (1999). Human KRML (MAFB): cDNA cloning, genomic structure, and evaluation as a candidate tumor suppressor gene in myeloid leukemias. *Genomics* **59**: 275–281.

Watson SK, deLeeuw RJ, Horsman DE, Squire JA, Lam WL. (2007). Cytogenetically balanced translocations are associated with focal copy number alterations. *Hum Genet* **120**: 795–805.

Watson SK, deLeeuw RJ, Ishkanian AS, Malloff CA, Lam WL. (2004). Methods for high throughput validation of amplified fragment pools of BAC DNA for constructing high resolution CGH arrays. *BMC Genomics* **5**: 6.

Weinstein IB. (2002). Cancer. Addiction to oncogenes—the Achilles heal of cancer. *Science* **297**: 63–64.

Wolf M, Mousses S, Hautaniemi S, Karhu R, Huusko P, Allinen M *et al.* (2004). High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. *Neoplasia* **6**: 240–247.

Zhou BB, Peyton M, He B, Liu C, Girard L, Caudler E *et al.* (2006). Targeting ADAM-mediated ligand cleavage to inhibit HER3 and EGFR pathways in non-small cell lung cancer. *Cancer Cell* **10**: 39–50.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).

Loss and Reduction of Fus1 Protein Expression is a Frequent Phenomenon in the Pathogenesis of Lung Cancer

Ludmila Prudkin,[1] Carmen Behrens,[2] Diane D. Liu,[3] Xian Zhou,[3] Natalie C. Ozburn,[2] B. Nebiyou Bekele,[3] John D. Minna,[5,6,7] Cesar Moran,[1,2] Jack A. Roth,[4] Lin Ji,[4] and Ignacio I. Wistuba[1,2]

Abstract **Purpose:** *FUS1*, a novel tumor-suppressor gene located in the chromosome 3p21.3 region, may play an important role in lung cancer development. Currently, *FUS1*-expressing nanoparticles have been developed for treating patients with lung cancer. However, the expression of Fus1 protein has not been examined in a large series of lung cancers and their sequential preneoplastic lesions.

Experimental Design: Using tissue microarrays, we examined Fus1 immunohistochemical expression in 281 non–small cell lung carcinoma (NSCLC) and 22 small cell lung carcinoma tissue specimens and correlated the findings with patients' clinicopathologic features. To investigate the expression of Fus1 in the early sequential pathogenesis of NSCLC, we studied Fus1 expression in 211 histologically normal and mildly abnormal bronchial epithelia, and 118 bronchial and alveolar preneoplastic lesions obtained from patients with lung cancer.

Results: Loss and reduction of expression was detected in 82% of NSCLCs and 100% of small cell lung carcinomas. In NSCLCs, loss of Fus1 immunohistochemical expression was associated with significantly worse overall survival. Bronchial squamous metaplastic and dysplastic lesions expressed significantly lower levels of Fus1 compared with normal ($P = 0.014$ and 0.047, respectively) and hyperplastic ($P = 0.013$ and 0.028, respectively) epithelia.

Conclusions: Our findings show a high frequency of Fus1 protein loss and reduction of expression in lung cancer, and suggests that this reduction may play an important role in the early pathogenesis of lung squamous cell carcinoma. These findings support the concept that *FUS1* gene and Fus1 protein abnormalities could be used to develop new strategies for molecular cancer therapy for a significant subset of lung tumors.

Lung cancer is the leading cause of cancer-related deaths in the United States (1). Lung cancer consists of several histologic types (2), the most frequent being small cell lung carcinoma (SCLC) and two types of non–small cell lung carcinoma (NSCLC), adenocarcinoma, and squamous cell carcinoma (3). In spite of advances, the underlying processes involved in the early pathogenesis of lung cancer remain unclear. NSCLCs are believed to arise after the progression of sequential preneoplastic lesions, including bronchial squamous dysplasias for squamous cell carcinoma and atypical adenomatous hyperplasias (AAH) for a subset of adenocarcinomas (4). An increased understanding of the molecular mechanisms involved in the pathogenesis and progression of lung cancer may lead to new and more effective strategies for early detection and targeted chemoprevention and treatment.

Tumor-suppressor genes play a major role in the pathogenesis of human lung and other cancers (2). Lung cancer cells harbor mutations and deletions in multiple known tumor-suppressor genes; however, genetic alterations and allelic losses (loss of heterozygosity) on the short arm of chromosome 3 sites (3p25–26, 3p21.3–22, 3p14, and 3p12) are among the most frequent and earliest molecular abnormalities detected in the pathogenesis of lung cancer (5, 6). In particular, chromosomal abnormalities at the 3p21.3 region and expressional deficiencies in 3p21.3 genes are frequently found in lung cancer (7). In addition, 3p21.3 allelic losses have been frequently detected in histologically normal bronchial epithelia and preneoplastic lesions in lung cancer patients and smokers (6, 8).

The novel *FUS1* gene is one of the candidate tumor-suppressor genes that have been identified in a 120-kb homozygous deletion region in human chromosome 3p21.3 (5, 9–11). Genomic alterations of the *FUS1* gene and resultant loss of expression or deficiency of posttranslational modification of

---

Authors' Affiliations: Departments of [1]Pathology, [2]Thoracic/Head and Neck Medical Oncology, [3]Biostatistics and Applied Mathematics, and [4]Thoracic Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas and [5]Hamon Center for Therapeutic Oncology Research, and Departments of [6]Internal Medicine and [7]Pharmacology, The University of Texas Southwestern Medical Center, Dallas, Texas

Received 5/22/07; revised 7/27/07; accepted 10/10/07.

Grant support: Specialized Program of Research Excellence in Lung Cancer (grant P50CA70907), National Cancer Institute, Bethesda, MD, and Department of Defense (grant W81XWH-04-1-0142).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Requests for reprints: Ignacio I. Wistuba, Department of Pathology, Unit 85, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030-4009. Phone: 713-563-9184; Fax: 713-792-0309; E-mail: iiwistuba@mdanderson.org.

©2008 American Association for Cancer Research.

doi:10.1158/1078-0432.CCR-07-1252

the Fus1 protein have been found in a majority of NSCLC cell lines and in almost all SCLCs (9 – 11). Recently, it was reported that Fus1 is a myristoylated protein and that myristoylation in its NH_2 terminus is required for *FUS1*-mediated tumor suppression activity (9). Immunohistochemical Fus1 expression examination of 20 NSCLC tissue specimens showed loss of protein in 15 of 20 (75%) cases, and these findings were confirmed by mass spectrometric analysis (11). To translate these findings into clinical applications for molecular cancer therapy, a novel *FUS1*-expressing nanoparticle has been developed for treating patients with lung cancer (12), suggesting that *FUS1* gene and protein abnormalities could be used to develop new strategies for molecular cancer therapy. To date, however, the expression of Fus1 has not been studied comprehensively in lung cancer tumors and lung preneoplastic lesion tissues.

To better understand the importance of Fus1 expression in lung cancer pathogenesis and progression, we investigated Fus1 immunohistochemical expression in a large series of NSCLC and SCLC tumor tissue specimens and adjacent lung bronchial and alveolar epithelial foci using tissue microarray specimens, and we correlated those findings with the clinicopathologic features of patients with lung cancer.

Materials and Methods

Case selection and tissue microarray construction. We obtained archival, formalin-fixed and paraffin-embedded material from surgically resected lung cancer specimens containing tumor and adjacent lung tissues from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M.D. Anderson Cancer Center (Houston, TX) from 1997 to 2001. This study was approved by the Institutional Review Board. Tumor tissue specimens from 303 lung cancers (22 SCLCs, 172 adenocarcinomas, and 109 squamous cell carcinomas) were histologically examined, classified using the 2004 WHO classification system (3), and selected for tissue microarray construction. After histologic examination, the tissue microarrays were constructed using triplicate 1 mm diameter cores from each tumor.

Detailed clinical and pathologic information, including demographic data, smoking history (never- and ever-smokers) and status (never, former, and current smokers), pathologic tumor-node-metastasis staging (13), overall survival, and time of recurrence, was available in most cases (Table 1). Patients who had smoked at least 100 cigarettes in their lifetime were defined as smokers, and smokers who quit smoking at least 12 months before lung cancer diagnosis were defined as former smokers.

To assess Fus1 immunohistochemical expression in the pathogenesis of NSCLC from the surgically resected formalin-fixed and paraffin-embedded specimens, we included 329 bronchial epithelium specimens which had normal histology (n = 68), basal cell hyperplasia (n = 120), squamous metaplasia (n = 23), squamous dysplasia (n = 62), and AAH lesions (n = 56) for analysis (Table 1). Histologic classification of epithelial lesions was done using the 2004 WHO classification system for lung preneoplastic lesions (3). For Fus1 expression analysis, squamous dysplasias were arranged into two groups: (*a*) low-grade, mild and moderate dysplasias (n = 14); and (*b*) high-grade, severe dysplasia and carcinoma *in situ* (n = 48).

Immunohistochemical staining and evaluation. The rabbit anti-Fus1 polyclonal antibody used for immunohistochemical staining was raised against a synthetic oligopeptide derived from NH_2-terminal amino acid sequence (NH_2-GASGSKARGLWPFAAC; ref. 11). Formalin-fixed and paraffin-embedded tissue histology sections (5 μm thick) were deparaffinized, hydrated, and heated in a steamer for 10 min with 10 mmol/L of sodium citrate (pH 6.0) for antigen retrieval. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% bovine serum albumin in TBS-t for 30 min. The slides were incubated with primary antibody at 1:400 dilution for 65 min at room temperature. After washing with PBS, incubation with biotin-labeled secondary antibody for 30 min followed. Finally, the samples were incubated with a 1:40 solution of streptavidin-peroxidase for 30 min. The staining was then developed with 0.05% 3′,3-diaminobenzidine tetrahydrochloride prepared in 0.05 mol/L of Tris buffer at pH 7.6 containing 0.024% H_2O_2 and then counterstained with hematoxylin. Formalin-fixed and paraffin-embedded lung tissues with normal bronchial epithelia were used as a positive control. For a negative control, we used the same specimens used for the positive controls, replacing the primary antibody with PBS.

Fus1 immunostaining was detected in the cytoplasm of epithelial and tumor cells (Fig. 1). Immunohistochemical expression was quantified by two independent pathologists (L. Prudkin and I.I. Wistuba) using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage of the reactivity extent. A consensus value on both intensity and extension was reached by the two independent observers.

Table 1. Demographic and clinicopathologic data regarding the lung cancer and respiratory epithelial samples studied for Fus1 immunohistochemical expression

Type and histology of samples	No.	Sex		Stage*				Smoking history [†]		Smoking status [‡]		
		F	M	I	II	III	IV	Yes	No	Never	Former	Current
Total cancers	303	155	148	—	—	—	—	—	—	—	—	—
SCLC	22	7	15	—	—	—	—	22	0	0	3	7
NSCLC	281	148	133	178	57	38	8	194	80	80	117	73
Adenocarcinoma	172	105	67	118	22	27	5	108	61	61	63	41
Squamous cell carcinoma	109	43	66	60	35	11	3	86	19	19	54	32
Total epithelial foci	329	136	193	—	—	—	—	274	50	50	120	145
Normal epithelium	68	35	33	—	—	—	—	53	15	15	26	27
Hyperplasia	120	31	89	—	—	—	—	107	11	11	50	57
Squamous metaplasia	23	10	13	—	—	—	—	18	2	2	9	9
Low grade dysplasia	14	6	8	—	—	—	—	14	0	0	8	6
High grade dysplasia	48	12	36	—	—	—	—	44	4	4	11	24
AAH	56	42	14	—	—	—	—	38	18	18	16	22

*Staging is shown only for NSCLC cases. All SCLCs were limited stage.

[†] Patient smoking history was not available for seven NSCLC and five epithelial specimens.

[‡] Patient smoking status was not available for 12 SCLC, 11 NSCLC, and 14 epithelial specimens.

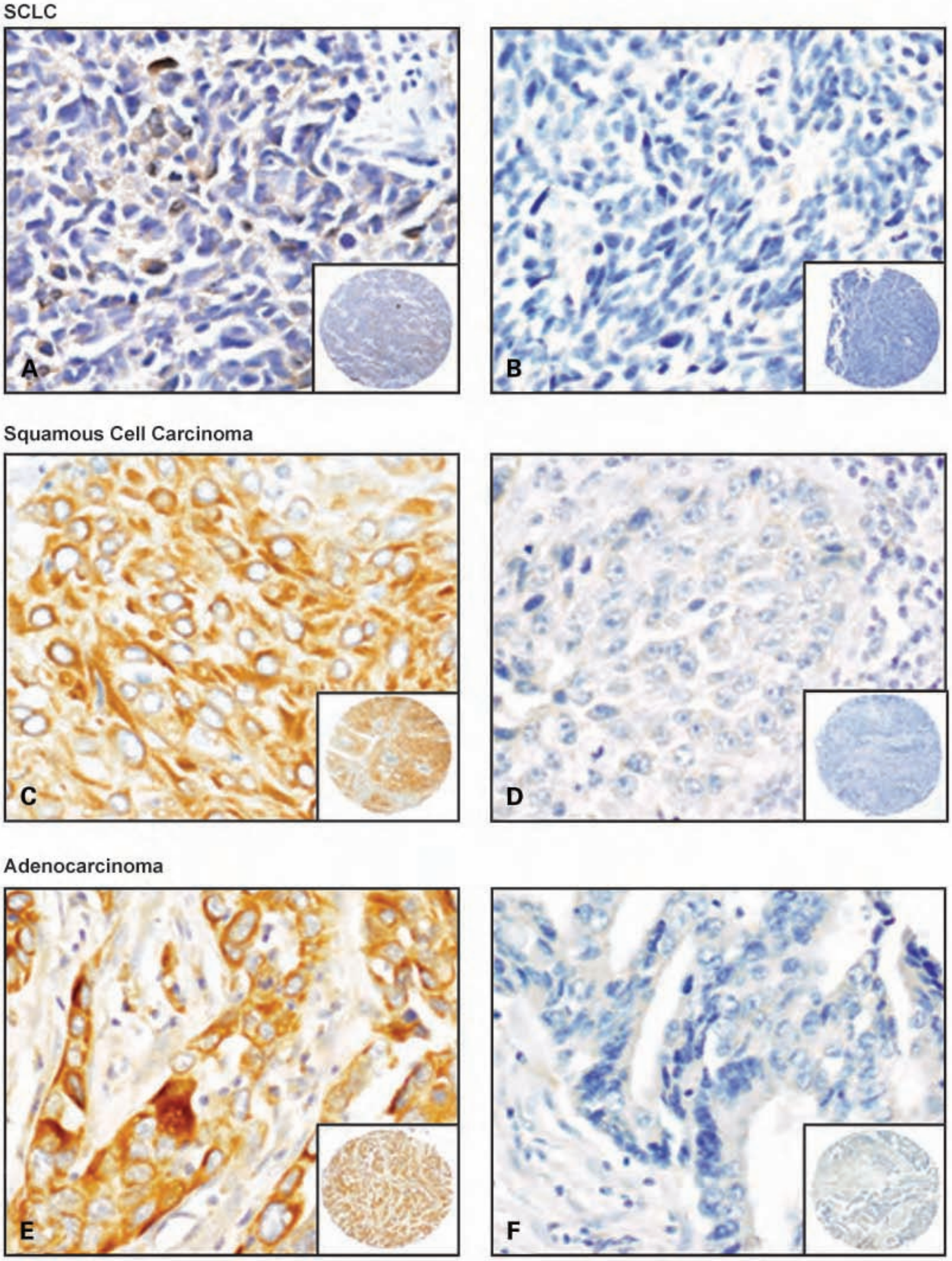


Fig. 1. Representative examples of Fus1 immunohistochemical staining of lung cancer specimens. SCLC with reduced (*A*) and negative (*B*) Fus1 expression. Squamous cell carcinoma with high (*C*) and negative (*D*) Fus1 expression. Adenocarcinoma with high (*E*) and negative (*F*) Fus1 expression. *A* to *F*, original magnification, ×400 (*pictures*), and ×40 (*insets*).

Correlation analyses were done between the three quantifications and all showed statistical significance ($r = 0.75$; $P < 0.001$). A final consensual score was obtained by multiplying both intensity and extension values (range, 0 – 300), and four levels of staining were calculated based on that score: (*a*) negative (score 0), (*b*) low (score ≤100), (*c*) intermediate (score >100 to ≤200), and (*d*) and high (score >200) expressions. On the basis of the high level of expression detected in normal bronchial epithelium specimens, high score levels were defined as preserved staining pattern, whereas intermediate and low score levels were defined as reduced staining pattern, and negative score as loss of expression. Levels and scores were used for analysis.

Statistical analysis. The data were summarized using standard descriptive statistics and frequency tabulations. Associations between categorical variables were assessed via cross-tabulation, χ^2 test, and Fisher's exact test. Wilcoxon rank sum test and Kruskal-Wallis test were done to assess the differences between patients' clinicopathologic groups with respect to continuous variables. Survival curves were estimated using the Kaplan-Meier method. Univariate Cox proportional hazard models were used to assess the effect of covariates on overall survival and recurrence-free survival. All computations were done using SAS and Splus 2000 (Insightful) statistical software. The mixed effect model was used to assess the differences in scores between normal and abnormal epithelia. The generalized estimating equation approach was used to estimate differences in the means for the data in Table 2.

Results

Fus1 immunohistochemical expression in lung cancer specimens. Lung cancer histologies varied in their pattern of immunohistochemical expression of Fus1 in the cytoplasm of the tumor cells. We detected a statistically significant difference ($P = 0.001$) in the Fus1 expression mean score among the three major types of lung cancer histologies examined. SCLCs had the lowest mean score (57; SD, 67.4), with tumors having either no protein expression (41%) or reduced expression (59%; Table 2; Fig. 1). Squamous cell carcinomas and adenocarcinomas showed intermediate levels of Fus1 expression, with mean scores of 127 (SD, 91.8) and 111 (SD, 79.1), respectively. Overall, 82% (230 of 281) of NSCLCs had lower Fus1 expressions (69%) or no Fus1 expression (13%). We note that the difference between squamous cell and adenocarcinoma histologies was not shown to be statistically significant at the 0.05 level ($P = 0.07$; Table 2). There was a significant difference ($P = 0.0008$) in the levels of Fus1 expression between SCLCs and NSCLCs (Table 2). Overall, lung tumor specimens showed lower scores and levels of expression compared with normal epithelium. Most tumors (231 out of 303, 76%; 91% of SCLCs and 75% of NSCLCs) had a score of <200, which was never associated with normal tissue.

Correlation between Fus1 immunohistochemical expression in NSCLC and clinicopathologic features. Using Fus1 expression mean scores and score levels, we detected no statistically significant correlation and/or association between protein expression and the clinicopathologic data, including sex, age, ethnicity, smoking history, and tumor-node-metastasis pathologic stage. Overall and disease-free survival analyses for Fus1 expression were done in 280 patients with tumor-node-metastasis stages I to IV (median follow-up, 3.90 years), and in 218 patients with stages I and II (median follow-up, 4.03 years) NSCLCs, who did not receive neoadjuvant or adjuvant therapies. Of interest, in both univariate and multivariate Cox model analyses, the hazard ratios for overall survival were much lower in cases having any level of Fus1 expression (low, intermediate, and high) compared with absence (negative) of protein expression (Table 3). These differences were statistically significant in most comparisons. Although the hazard ratios for recurrence-free survival showed similar trends than overall survival, the P values were not statistically significant (Supplementary Table S1).

Fus1 immunohistochemical expression in the sequential pathogenesis of lung cancer. To characterize the pattern of Fus1 expression in the sequential pathogenesis of NSCLC, we investigated the protein immunohistochemical expression in histologically normal epithelium, hyperplasia, squamous metaplasia, and squamous dysplasia bronchial sites obtained from surgically resected NSCLC specimens (Fig. 2). We also examined 56 AAH lesions, a putative precursor lesion for adenocarcinomas (3, 4). Mean scores and score levels of Fus1 cytoplasmic expression in the epithelia were used for comparison between all different epithelial histologic categories. Overall, normal, mildly abnormal, and preneoplastic respiratory epithelia showed higher mean scores and score

Table 2. Fus1 immunohistochemical expression in lung cancer and epithelial foci using tissue microarray specimens

Histology of samples	No. of samples	Fus1 score, mean (SD)	Fus1 score levels				
			Lost (negative) *n* (%)	Reduced (low + intermediate) *n* (%)	Preserved (high) *n* (%)	*P* value, Fus1 levels	
Cancer specimens						Comparison between tumors	
SCLC	22	57 (67.4)	9 (41)	13 (59)	0	0.0008	
NSCLC	281	121 (87.3)	36 (13)	194 (69)	51 (18)		
Adenocarcinoma	172	127 (91.8)	25 (15)	110 (64)	37 (22)	0.07	
Squamous cell carcinoma	109	111 (79.1)	11 (10)	84 (77)	14 (13)		
Epithelial specimens						Comparison between normal and abnormal epithelia*	
						Score	Level
Normal epithelium	68	251 (46.6)	0	27 (40)	41 (60)	Ref [†]	Ref [†]
Basal cell hyperplasia	120	263 (47.5)	0	40 (33)	80 (67)	0.35	0.68
Squamous metaplasia	23	215 (46.0)	0	18 (78)	5 (22)	0.014	0.014
Squamous dysplasia	62	224 (51.3)	0	25 (45)	31 (55)	0.0004	0.047
Low-grade dysplasia	14	200 (39)	0	12 (86)	2 (14)	0.010	0.28
High-grade dysplasia	48	231 (53)	0	23 (48)	25 (52)	0.005	0.019
AAH	56	250 (52.3)	0	25 (45)	31 (55)	0.61	0.79

*Mixed effect model was used to assess difference of scores between normal and abnormal epithelia. A GEE model was used to compare the differences in FUS1 level between normal and abnormal epithelia.

[†] Normal epithelium was used as reference value (Ref).

Table 3. Univariate and multivariate Cox model assessing effects of covariates on overall survival

(A) Univariate Cox model

Variable		All stages (n = 280)		Stages I and II (n = 213)	
		Hazard ratio	P	Hazard ratio	P
Age		1.044	<0.0001	1.057	<0.0001
Histology	Squamous cell vs. adenocarcinoma	1.585	0.02	1.667	0.03
Tumor*	T_{2+3+4} vs. T_1	2.024	0.002	—	—
	T_{2+3} vs. T_1	—	—	2.175	0.002
Lymph node*	N_{1+2} vs. N_0	2.122	0.0003	2.531	0.0002
	N_1 vs. N_0	—	—	—	—
Tumor stage*	II + III + IV vs. I	2.01	0.0005	—	—
	II vs. I	—	—	2.47	0.0002
Fus1 score mean		0.999	0.70	0.999	0.61
Fus1 score level	Low † vs. negative ‡	0.535	0.06	0.502	0.07
	Intermediate † vs. negative ‡	0.744	0.35	0.831	0.60
	High§ vs. negative ‡	0.685	0.28	0.536	0.13

(B) Multivariate Cox model

Variable		All stages		Stages I and II	
		Hazard ratio	P	Hazard ratio	P
Age		1.047	<0.0001	1.059	<0.0001
Histology	Squamous cell vs. adenocarcinoma	1.528	0.04	1.438	0.13
Tumor stage*	II + III + IV vs. I	2.163	0.0002	—	—
	II vs. I	—	—	2.519	0.0004
Fus1 score level	Low † vs. negative ‡	0.357	0.002	0.312	0.003
	Intermediate † vs. negative ‡	0.474	0.02	0.544	0.10
	High§ vs. negative ‡	0.518	0.06	0.415	0.03

*Pathologic tumor, lymph node, and metastasis stage.
†Low and intermediate, reduced Fus1 immunohistochemical expression.
‡Negative, absence of Fus1 immunohistochemical expression.
§High, preserved or normal Fus1 immunohistochemical expression.

levels of Fus1 expression compared with tumors, and none of the nonmalignant epithelial specimens showed loss of expression. Normal and hyperplastic epithelia displayed high mean scores and score levels of Fus1 expression (Table 2), with high expression score levels in 60% (41 of 68) and 67% (80 of 120) of normal and hyperplastic foci, respectively. Squamous metaplasia and dysplasia lesions had significantly lower score means and score levels of Fus1 expression compared with histologically normal and hyperplastic epithelia, with 78% and 57% of the squamous metaplastic and dysplastic lesions, respectively, showing a lower protein expression (Table 2). No differences were detected in the level of Fus1 expression comparing low-grade (mild and moderate dysplasias) and high-grade (carcinoma *in situ* and severe dysplasia) squamous dysplastic lesions. No significant differences in the mean scores and score levels of Fus1 expression were detected comparing normal and hyperplastic bronchial epithelium with AAH lesions. No significant associations were observed between Fus1 expression and age, sex, or smoking status of NSCLCs patients from whom epithelial specimens were obtained.

Discussion

Mutations in 3p21.3 genes are rarely found in human tumors, including lung cancer (7). Therefore, some mechanisms other than the classic two-hit model, which requires mutation in one allele and silencing or loss on another allele, might be of importance in the ultimate inactivation of 3p21.3 genes. These alternative mechanisms include promoter methylation, haploinsufficiency, altered RNA splicing, as well as defects in transcriptional, translational, and posttranslational processes. In lung cancer, only a few *FUS1* mutations that alter or truncate amino acid sequences have been detected (7), and its promoter methylation is a rare phenomenon (11). It has been hypothesized that this gene is inactivated in lung tumors by alternative mechanisms, such as influence from the stochastic effects of 3p21.3 allele haploinsufficiency (6) and a posttranslational modification of the gene product by deficient *N*-myristoylation of the Fus1 protein (11). It has been shown that myristoylation is required for *FUS1*-mediated tumor-suppressing activity, suggesting a novel mechanism for the inactivation of tumor-suppressor genes for human cancers (11). Uno et al. (11), using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, analyzed the *N*-myristoylation status of Fus1 protein in frozen lung cancer tissue specimens and determined that protein loss of immunohistochemical expression correlated with loss of protein myristoylation.

Thus, to investigate the frequency and pattern of Fus1 protein expression abnormalities in lung cancer, we examined the immunohistochemical expression of the protein using tissue microarrays in a large series of primary tumor specimens with annotated clinical information. Our findings confirmed and further expanded previously reported data on loss of Fus1 protein expression in 20 NSCLC tissue specimens (11). Lung tumors showed lower levels of expression than normal bronchial epithelium. We found that loss or reduction of

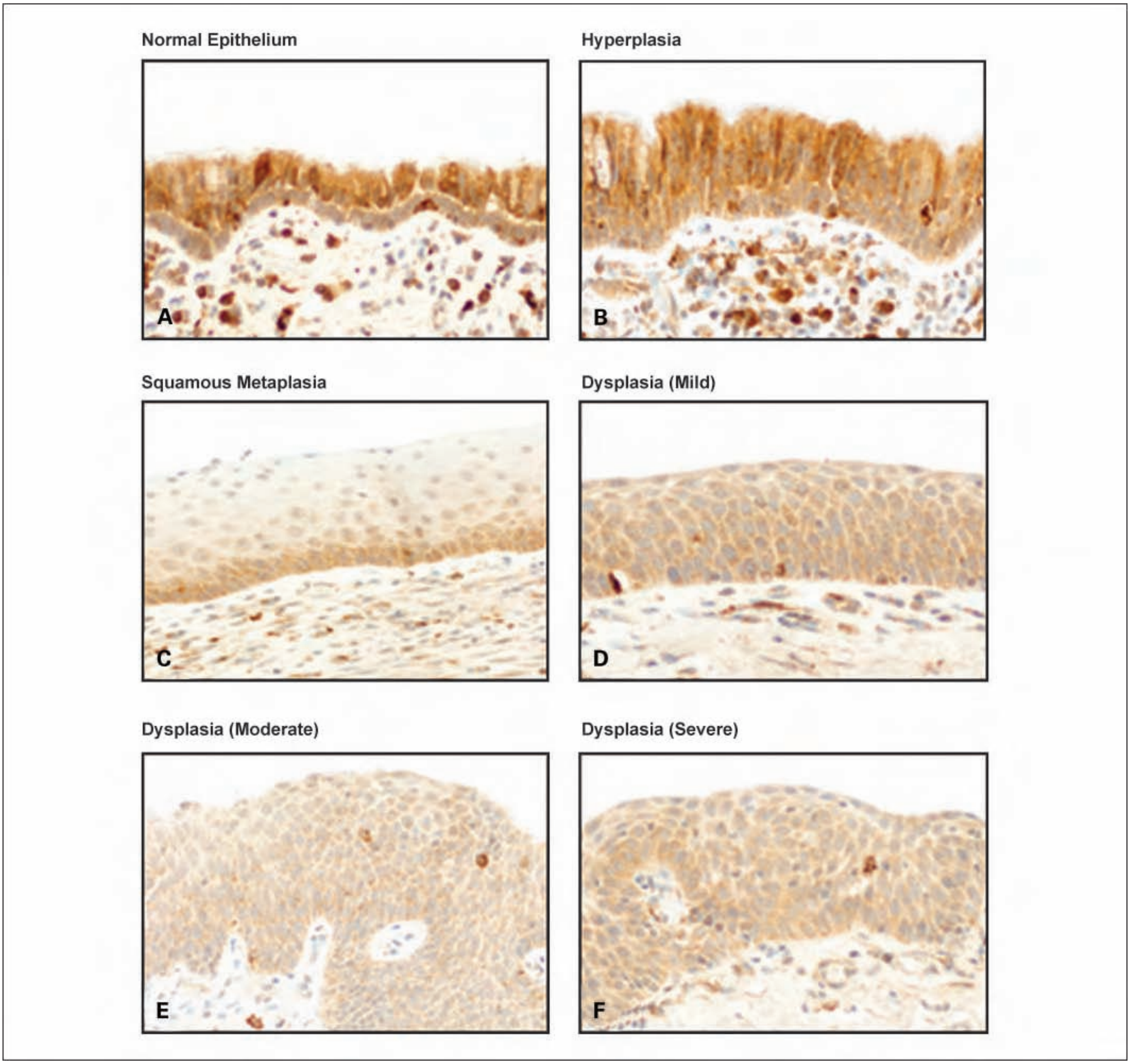


Fig. 2. Representative examples of Fus1 immunohistochemical expression in histologically normal (*A*), hyperplastic (*B*), metaplastic (*C*), and dysplastic (*D* – *F*) bronchial epithelia. High levels of expression are shown in the normal and hyperplastic epithelia and reduced levels in squamous metaplasia and dysplasias. Original magnification, ×400.

Fus1 immunohistochemical expression was present in all SCLCs and most (82%) of the NSCLCs (87% of the squamous cell carcinomas and 79% of the adenocarcinomas). Fus1 protein expression was absent in half of the SCLCs and in 13% of the NSCLCs. For NSCLCs, we found that loss of Fus1 expression is a significant independent adverse prognostic factor for patients' overall survival. Our findings that retention of any level of Fus1 immunohistochemical expression significantly correlates with better outcome in NSCLC are in agreement with the potent effect of tumor suppressor activity of *FUS1* gene as shown by *in vitro* and *in vivo* studies (11, 12). Although *FUS1* is one of the nine putative tumor-suppressor genes located in the 3p21.3 region frequently deleted in lung and other tumor sites (7), our study is the first report on Fus1 protein expression abnormalities in a large series of human tumors arising at any site.

The ability to rescue the tumor phenotype by gene inactivation after the replacement of those genes is one of the criteria needed for a gene to be considered a tumor-suppressor gene. There is evidence indicating that the replacement of *FUS1* in nonexpressive NSCLC cell lines inhibits cell growth and induces apoptosis (9, 10). More importantly, Ji et al. (9) showed that intratumoral injection of adenovirus-*FUS1* significantly reduces tumor growth in *FUS1* region–deficient tumor xenografts and experimental metastasis. Following *in vitro* results, Ito et al. (12) reported that intratumoral and intravenous administration of *FUS1* complexed to nanoparticles in mice bearing human lung cancer cell line xenografts resulted in the inhibition of tumor growth, a decreased number of metastases, and prolonged survival compared with results for untreated animals. The restoration of tumor-suppressor genes altered in cancer development is currently a valid therapeutic approach. On the basis of preclinical studies, a phase I clinical trial by our group using *FUS1*-mediated molecular therapy by systemic administration of *FUS1*-nanoparticles is currently under way in stage IV NSCLC patients at The University of Texas M.D. Anderson Cancer Center. Our findings of a lack or reduction of expression of Fus1 protein in most lung tumors supports the concept of the delivery of functional *FUS1* as an effective therapeutic strategy for human lung cancer.

Lung cancers are believed to be the consequence of a series of progressive preneoplastic changes in the respiratory mucosa that accumulate a sequence of genetic changes (14). These genetic abnormalities are frequently extensive and multifocal throughout the respiratory epithelium, indicating a field effect or field cancerization phenomenon (4). Although the sequential preneoplastic changes have been defined for squamous carcinomas, they have been poorly documented for lung adenocarcinomas and SCLCs (4). Mucosal changes in the large airways that may precede squamous cell carcinoma include squamous dysplasia and carcinoma *in situ* in the central airway (15, 16). Adenocarcinomas may be preceded by morphologic changes, including AAH in peripheral airway cells (17). In squamous cell carcinoma pathogenesis, genetic abnormalities commence in histologically normal epithelium and augment with increasing severity of histologic changes, with 3p21.3 allelic loss as the earliest genetic change being detected in patients with lung cancer and smokers (6, 8). Our findings of a significant reduction of Fus1 protein immunohistochemical expression in squamous metaplasia and dysplasia histologies compared with histologically normal and hyperplastic epithelia suggest that the reduction and partial inactivation of 3p21.3 *FUS1* gene expression is an early phenomenon in the pathogenesis of squamous cell carcinoma. AAH lesions showed some reduction in the level of Fus1 immunohistochemical expression in a subset of cases (45%), but no differences compared with histologically normal and hyperplastic epithelia were detected. Although 3p12 and 3p14 allelic losses have been shown in ~20% of AAHs, no detailed mapping analysis of chromosome 3p that includes the *FUS1* 3p21.3 region has been done in these lesions (18).

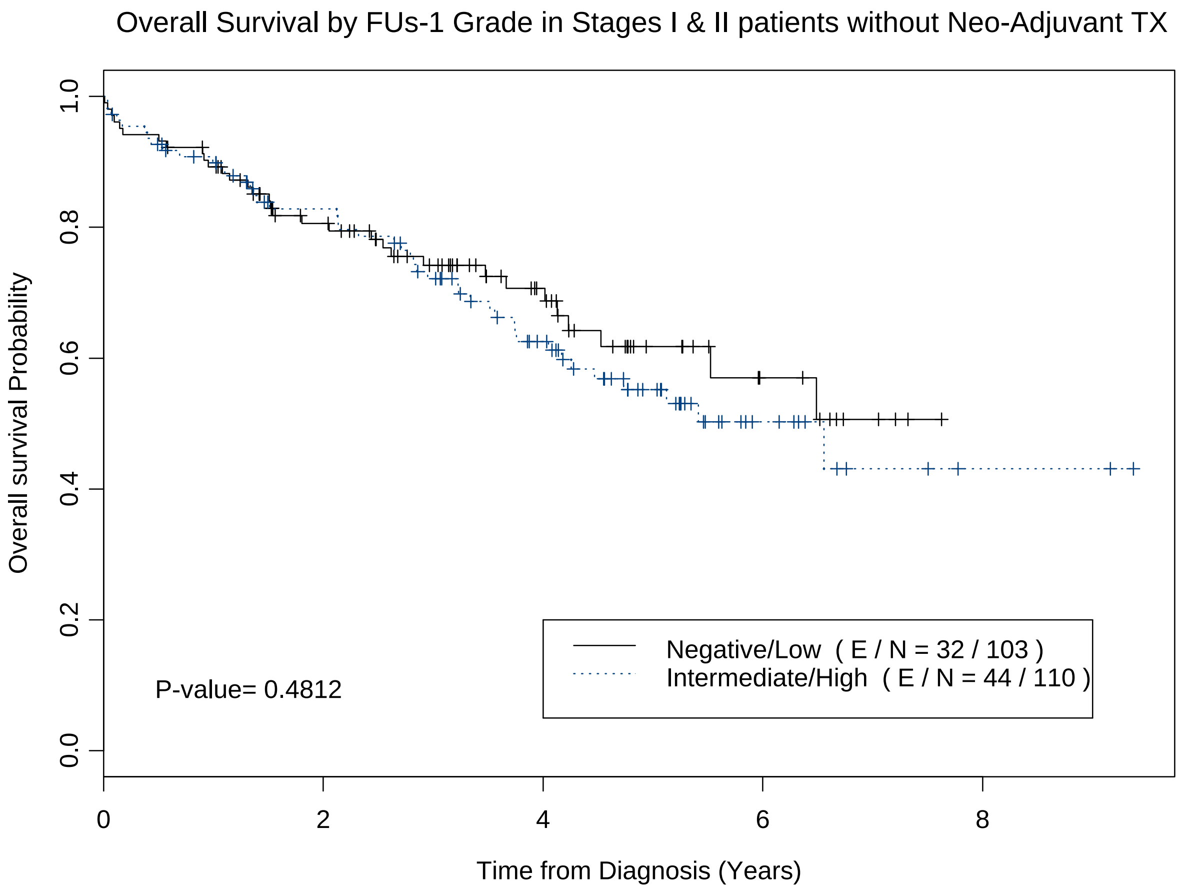
Although a total loss of Fus1 protein immunohistochemical expression was detected in 50% of SCLCs and in 13% of NSCLSs, no normal and abnormal epithelial sites showed a complete lack of Fus1 protein expression. Because most lung cancers and adjacent preneoplastic lesions have shown allelic loss at the 3p21.3 region (6), *FUS1* haploinsufficiency may play a role in the inactivation of Fus1 protein in lung cancer pathogenesis (19). In diploid cells, each gene exists in two copies, in contrast to haploids in which each cell contains a single copy of the genome. When one of the alleles is mutated or deleted, there is an ~50% reduction in the level of proteins synthesized. Generally, the haploinsufficiency occurs when the level of proteins synthesized decreases below a threshold level and is insufficient for the onset of some desired biological activity. We have detected lower Fus1 protein expression in a subset of all epithelial foci examined, including histologically normal, hyperplastic, metaplastic, and preneoplastic epithelia. Interestingly, the reduction of Fus1 expression in epithelial specimens, including 40% of histologically normal epithelia, was observed in samples from both smoker and never-smoker NSCLC patients, suggesting that this phenomenon is not necessarily associated with smoking. We speculate that allelic loss and the haplotype in the *FUS1* 3p21.3 region at very early stages of lung tumor pathogenesis may lead to a reduction of Fus1 protein synthesis. We hypothesize that a deficiency in myristoylation might lead to a greater reduction and complete loss of Fus1 protein expression in the later stages of tumor development, such as in microinvasive and invasive lesions.

In summary, our findings show a high frequency of Fus1 protein reduction and loss of expression in SCLC and NSCLC tissue specimens, and suggest that reduction of this protein may play an important role in the early pathogenesis of squamous cell carcinomas. All these findings support the concept that *FUS1* gene and protein abnormalities could be used to develop new strategies for molecular cancer therapy for a significant subset of lung tumors.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. CA Cancer J Clin 2006;56:106–30.
2. Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. Cancer Cell 2002;1:49–52.
3. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Pathology and genetics: tumours of the lung, pleura and heart, editors. International Agency for Research on Cancer (IARC). Lyon: IARC; 2004.
4. Wistuba I. Genetics of preneoplasia: lessons from lung cancer. Curr Mol Med 2007;7:3–14.
5. Lerman MI, Minna JD. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium. Cancer Res 2000;60: 6116–33.
6. Wistuba II, Behrens C, Virmani AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.
7. Ji L, Minna JD, Roth JA. 3p21.3 tumor suppressor cluster: prospects for translational applications. Future Oncol 2005;1:79–92.
8. Wistuba II, Lam S, Behrens C, et al. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89:1366–73.
9. Ji L, Nishizaki M, Gao B, et al. Expression of several genes in the human chromosome 3p21.3 homozygous deletion region by an adenovirus vector results in tumor suppressor activities *in vitro* and *in vivo*. Cancer Res 2002;62:2715–20.
10. Kondo M, Ji L, Kamibayashi C, et al. Overexpression of candidate tumor suppressor gene FUS1 isolated from the 3p21.3 homozygous deletion region leads to G1 arrest and growth inhibition of lung cancer cells. Oncogene 2001;20:6258–62.
11. Uno F, Sasaki J, Nishizaki M, et al. Myristoylation of the fus1 protein is required for tumor suppression in human lung cancer cells. Cancer Res 2004;64: 2969–76.
12. Ito I, Ji L, Tanaka F, et al. Liposomal vector mediated delivery of the 3p FUS1 gene demonstrates potent antitumor activity against human lung cancer *in vivo*. Cancer Gene Ther 2004;11:733–9.
13. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111:1710–7.
14. Wistuba II, Mao L, Gazdar AF. Smoking molecular damage in bronchial epithelium. Oncogene 2002;21: 7298–306.
15. Colby TV, Wistuba II, Gazdar A. Precursors to pulmonary neoplasia. Adv Anat Pathol 1998;5:205–15.
16. Hastings RH, Auger WR, Kerr KM, Quintana RA, Deftos LJ. Parathyroid hormone-related protein and lung injury after pulmonary thromboendarterectomy. Regul Pept 2001;102:1–7.
17. Westra WH, Slebos RJ, Offerhaus GJ, et al. K-ras oncogene activation in lung adenocarcinomas from former smokers. Evidence that K-ras mutations are an early and irreversible event in the development of adenocarcinoma of the lung. Cancer 1993;72:432–8.
18. Yamasaki M, Takeshima Y, Fujii S, et al. Correlation between genetic alterations and histopathological subtypes in bronchiolo-alveolar carcinoma and atypical adenomatous hyperplasia of the lung. Pathol Int 2000;50:778–85.
19. Zabarovsky ER, Lerman MI, Minna JD. Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. Oncogene 2002;21:6915–35.

Supplementary Figure S1. Overall Survival curve



Supplementary Table S1. Univariate Cox model assessing effects of covariates on recurrence free survival

Variable		All stages (N=280)		Stages I and II (N=213)	
		Hazard ratio	*p*-Value	Hazard ratio	*p*-Value
Age		1.012	0.30	1.028	0.08
Histology	Squamous cell vs. adenocarcinoma	0.876	0.61	1.185	0.5892
Tumor[1]	T2+3+4 vs. T1	1.404	0.19	--	--
	T2+3 vs. T1	--	--	1.352	0.3412
Lymph node[1]	N1+2 vs. N0	2.629	0.0001	--	--
	N1 vs. N0	--	--	1.852	0.092
Tumor stage[1]	II+III+IV vs. I	2.564	0.0001	--	--
	II vs. I	--	--	1.916	0.0596
Fus1 score mean		1.000	0.75	0.999	0.6923
Fus1 score level	Low[2] vs. negative[3]	0.645	0.29	0.551	0.23
	Intermediate[2] vs. negative[3]	0.790	0.57	0.708	0.48
	High[4] vs. negative[3]	0.939	0.89	0.614	0.37

[1] Pathology TNM stage = tumor, lymph node and metastasis.

[2] Low and intermediate = reduced Fus1 IHC expression.

[3] Negative = absence of Fus1 IHC expression.

[4] High = preserved or normal Fus1 IHC expression.

Supplementary Table S2. Demographic and clinico-pathologic data regarding the lung cancer and respiratory epithelial samples studied for Fus1 immunohistochemical expression

Type and histology of samples	No. (%)	Sex (%)		Stage[1] (%)				Smoking history[2] (%)		Smoking status[3] (%)		
		F	M	I	II	III	IV	Yes	No	Never	Former	Current
Total cancers	303	155 (51.2)	148 (48.8)	--	--	--	--	--	--	--	--	--
SCLC	22 (7.3)	7 (31.9)	15 (68.2)	--	--	--	--	22 (100)	0 (0)	0 (0)	3 (30)	7 (70)
NSCLC	281 (92.7)	148 (52.7)	133 (47.3)	178 (63.4)	57 (20.3)	38 (13.5)	8 (2.8)	194 (70.8)	80 (29.2)	80 (29.6)	117 (43.3)	73 (27.1)
Adenocarcinoma	172 (61.2)	105 (61)	67 (39)	118 (68.6)	22 (12.8)	27 (15.7)	5 (2.9)	108 (63.9)	61 (36.1)	61 (36.9)	63 (38.2)	41 (25.7)
Squamous cell carcinoma	109 (38.8)	43 (39.5)	66 (60.5)	60 (55)	35 (32.1)	11 (10.1)	3 (2.8)	86 (81.9)	19 (18.1)	19 (18.1)	54 (51.4)	32 (30.5)
Total epithelial foci	329	136 (41.3)	193 (58.7)	--	--	--	--	274 (84.6)	50 (15.4)	50 (15.9)	120 (38.1)	145 (46)
Normal epithelium	68 (20.6)	35 (51.5)	33 (48.5)	--	--	--	--	53 (77.9)	15 (22.1)	15 (22.1)	26 (38.2)	27 (38.7)
Hyperplasia	120 (36.5)	31 (25.8)	89 (74.2)	--	--	--	--	107 (90.7)	11 (9.3)	11 (9.3)	50 (42.4)	57 (48.3)
Squamous metaplasia	23 (7)	10 (43.5)	13 (56.5)	--	--	--	--	18 (90)	2 (10)	2 (10)	9 (45)	9 (45)
Low grade dysplasia	14 (4.3)	6 (42.9)	8 (57.1)	--	--	--	--	14 (100)	0 (0)	0 (0)	8 (57.1)	6 (42.9)
High grade dysplasia	48 (14.6)	12 (25)	36 (75)	--	--	--	--	44 (91.7)	4 (8.3)	4 (10.3)	11 (28.2)	24 (61.5)
AAH[4]	56 (17)	42 (75)	14 (25)	--	--	--	--	38 (67.8)	18 (32.2)	18 (32.2)	16 (28.6)	22 (39.2)

[1] Staging is shown only for NSCLC cases. All SCLCs were limited stage.
[2] Patient smoking history was not available for seven NSCLC and five epithelial specimens.
[3] Patient smoking status was not available for 12 SCLC, 11 NSCLC, and 14 epithelial specimens.
[4] AAH = atypical adenomatous hyperplasia.

Epithelial-Mesenchymal Transition in Non-Small Cell Lung Cancer Development and Progression

Ludmila Prudkin,[1] Diane Liu,[2] Natalie C. Ozburn,[3] Menghong Sun,[1] Carmen Behrens,[3] Ximing Tang,[3] Kathlynn C. Brown,[4,5] B. Nebiyou Bekele,[2] Cesar Moran,[1] and Ignacio I. Wistuba[1, 3]

From the Departments of [1]Pathology, [2]Biostatistics, and [3]Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA; and [4]Division of Translational Research, Department of Internal Medicine, and [5]Simmons Comprehensive Cancer Center, The University of Texas Southwestern Medical Center, Dallas, Texas.

Running title: EMT in NSCLC development and progression

Grant support: This study was supported in part by a grant from the Specialized Program of Research Excellence (SPORE) in Lung Cancer (grant P50 CA70907) and grant 1RO1CA106646, National Cancer Institute, Bethesda, MD, and Department of Defense (grant W81XWH-04-1-0142).

Requests for reprints:

Ignacio I. Wistuba

Department of Pathology, Unit 85

The University of Texas M. D. Anderson Cancer Center

1515 Holcombe Blvd., Houston, TX 77030-4009

tel: (713) 563-9184

fax: (713) 563-1848

e-mail: iiwistuba@mdanderson.org.

Abstract

Epithelial-to-mesenchymal transition (EMT) is a process in which cells undergo a developmental switch from an epithelial to a mesenchymal phenotype. We investigated the role of EMT in the pathogenesis and progression of non-small cell lung carcinoma (NSCLC). Archived tissue from adenocarcinoma and squamous cell carcinoma primary tumors (n=325) and brain metastases (n=48) and adjacent bronchial epithelial specimens (n=192) were analyzed for immunohistochemical expression by image analysis of E-cadherin, N-cadherin, integrin-αvβ6, vimentin, and matrix metalloproteinase-9. The findings were compared with patients' clinicopathologic features. High expression of the EMT phenotype (low E-cadherin and high N-cadherin, integrin-αvβ6, vimentin, and matrix metalloproteinase-9) was found in most NSCLC tumors. In primary tumors, the expression pattern varied according to the tumor histologic type. Low E-cadherin membrane and high N-cadherin cytoplasmic expression were significantly more common in squamous cell carcinoma than in adenocarcinoma (P=0.002 and 0.005, respectively). Dysplastic lesions had significantly lower expression of the EMT phenotype than did squamous cell carcinomas, and integrin-αvβ6 membrane expression increased stepwise according to the histopathologic severity. Brain metastases had decreased EMT expression compared with primary tumors. Brain metastases had significantly lower integrin-αvβ6 membrane (P=0.04) and N-cadherin membrane and cytoplasm (P<0.0002) expression than did primary tumors.

The EMT phenotype is commonly expressed in primary squamous cell carcinoma and adenocarcinoma of the lung; this expression occurs early in the pathogenesis of squamous cell carcinoma. NSCLC brain metastases showed characteristics of reversed mesenchymal-epithelial-transition. Our findings suggest that EMT is a potential target for lung cancer chemoprevention and therapy.

Key words: epithelial-to-mesenchymal transition, tissue microarray, immunohistochemical analysis, lung cancer, preneoplasia, brain metastases

Introduction

Lung cancer remains the leading cause of cancer-related death in the United States (1). It is a highly complex neoplasm (2) that includes several histologic types, the most common being small cell lung carcinoma (SCLC) and two types of non-SCLC (NSCLC), adenocarcinoma and squamous cell carcinoma (3). Despite therapeutic advances, the 5-year survival rate across all stages of the disease is approximately 15% 1, as many patients are not diagnosed until the disease is advanced. Surgery is considered the best treatment option, but only approximately 25% of NSCLC tumors are suitable for potentially curative resection (4). NSCLC is believed to result from a progressive series of lesions (5). This sequence has been clearly established in squamous cell carcinoma, but in adenocarcinoma, the progression sequence is less clear (5). A deeper understanding of the molecular mechanisms involved in the pathogenesis and progression of lung cancer may lead to new and more effective strategies for early detection and targeted chemoprevention and treatment.

Epithelial-to-mesenchymal transition (EMT) is a process in which cells undergo a developmental switch from an epithelial to a motile mesenchymal phenotype (6). This process has been related to embryologic morphogenesis, fibrosis, and lately, to the progression and metastasis of epithelial tumors (7). Solid tumor progression requires EMT for tumor cells to invade and metastasize. Colon, prostate, and breast cancer are tumors with the EMT phenotype (8). In lung cancer, EMT has been studied *in vitro*; the expression of individual markers in EMT has been described, and these markers are associated with prognosis (9, 10). These findings have led to the hypothesis that EMT is a target for lung cancer therapy. In addition, data suggest that EMT is influenced by epidermal growth factor receptor (EGFR) activity and predicts EGFR tyrosine kinase inhibitor (TKI) sensitivity (9, 11).

To better understand the importance of EMT in NSCLC pathogenesis and progression, we determined the immunohistochemical expression of five markers related to EMT (E-cadherin, N-cadherin, integrin-αvβ6, vimentin, and matrix metalloproteinase (MMP)-9) by image analysis of tissue microarray (TMA) specimens. We compared these findings with patients' clinicopathologic features and the immunohistochemical expression of EGFR and phosphorylated EGFR (pEGFR). In a subset of adenocarcinomas, we compared EMT expression with the presence of *EGFR* mutations. In addition, we investigated the role of EMT in the early pathogenesis of lung cancer by studying marker expression in the preneoplastic sequence of squamous cell carcinoma. Mesenchymal-to-epithelial transition (MET), the reverse of EMT, has been hypothesized to occur at the site of tumor metastasis (6, 8). Therefore, to determine whether MET was present in advanced lung tumors, we analyzed EMT markers in lung cancer brain metastases.

Materials and methods

Case selection and TMA construction. We obtained archived, formalin-fixed, paraffin-embedded (FFPE) tissue from surgically resected lung cancer specimens (lobectomies and pneumonectomies) containing tumor and adjacent normal and abnormal epithelium tissues from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX), which has been approved by the institutional review board. The tissue had been collected from 1997 to 2001. The tissue specimens were histologically examined and classified using the 2004 World Health Organization classification system (3). We selected 325 NSCLC tissue samples (209 adenocarcinomas and 116 squamous cell carcinomas) for our TMA. TMAs were constructed using triplicate 1-mm diameter cores per tumor; each core included central, intermediate, and peripheral tumor tissue. Detailed clinical and pathologic information, including demographics, smoking history (never- and ever-smokers), and smoking status (never, former, and current), clinical and pathologic TNM stage, overall survival duration, and time to recurrence, was available for most cases (Table 1). Tumors were pathologic TNM stages I-IV according to the revised International System for Staging Lung Cancer (12). The expression of integrin-αvβ6 in this NSCLC TMA has been previously described (13).

To assess the immunohistochemical expression of EMT markers in the early pathogenesis of NSCLC, we studied TMAs containing 192 bronchial epithelium specimens (normal histology, N=49; basal cell hyperplasia, N=76; squamous metaplasia, N=13; squamous dysplasia, N=34; and carcinoma *in situ*, N=20) from 89 patients with lung cancer. For statistical analysis, we grouped the lesions by tissue type: normal and reactive (including normal, hyperplasia, and squamous metaplasia) and dysplastic (including dysplasia and carcinoma *in situ*). Finally, to determine the expression of EMT markers in lung cancer metastasis, we examined TMAs containing 48 brain metastases (37 from adenocarcinomas and 11 from squamous cell carcinomas).

Immunohistochemical staining and evaluation. Antibodies against the following molecules were purchased and used: E-cadherin (Santa Cruz Biotechnology, Santa Cruz, CA; catalogue number Sc-8426; dilution 1:100), N-cadherin (Zymed, Carlsbad, CA; catalogue number 18-0224, concentration 1:100), integrin-αvβ6 (Calbiochem, Gibbstown, NJ; catalogue number 407317; dilution 1:300), vimentin (Dako, Carpinteria, CA; catalogue number M-0725; dilution 1:400), MMP-9 (Genetex, San Antonio, TX; catalogue number GTX-58899; dilution 1:400), EGFR, and pEGFR (Tyr-1086; Invitrogen, Carlsbad, CA; catalogue numbers 28-0005 and 36-9700, respectively; dilution 1:100). Immunohistochemical staining was performed as follows: 5-μM-thick FFPE tissue sections were deparaffinized, hydrated, heated in a steamer for 10 minutes with 10 mM sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was performed with 3% H_2O_2 in methanol at ambient temperature for 15 minutes, followed by 10% bovine serum albumin in tris-buffered saline-t at ambient temperature for 30 minutes. The slides were incubated with primary antibody at ambient temperature and washed with phosphate-buffered saline, followed by incubation with biotin-labeled secondary antibody for 30 minutes. Finally, the samples were incubated with a 1:40 solution of streptavidin-peroxidase for 30 minutes. Staining was developed with 0.05% 3', 3-diaminobenzidine tetrahydrochloride, which had been freshly prepared in 0.05 mol/L Tris buffer at pH 7.6 containing 0.024% H_2O_2 and then counterstained with hematoxylin, dehydrated, and mounted. FFPE NSCLC cell lines and tissues with normal bronchial epithelia were used as the positive control. As the negative control, we used the same specimens used as positive controls but replaced the primary antibody with phosphate-buffered saline.

Protein expression was quantified using the ARIOL image analysis system (Applied Imaging, San Jose, CA). This system gives an immunohistochemistry reading that avoids interobserver and intraobserver variability. For each marker, different cell sites were evaluated: membrane in E-cadherin and integrin-αvβ6, membrane and cytoplasm in N-cadherin, and cytoplasm in vimentin and MMP-9. In addition, EGFR and pEGFR (Tyr-1086) membranous staining was evaluated. Membrane expression was scored on the basis of the intensity and completeness of immunostaining, similar to Her2-Neu assessment classification (14). For the statistical analysis, we used class score, grouped into two categories: 0 and 1, low expression; and, 2 and 3, high expression. Cytoplasm staining was scored to account for both the intensity and extent of protein expression. These scores were used in the statistical analysis of cytoplasm expression. Vimentin is a special marker that is highly expressed in stromal and inflammatory cells, but it is not suitable for image analysis. Therefore, this marker was scored by a trained pathologist (L.P.) using light microscopy, a four-value intensity scale (0-3+), and percentage extent (0%-100%). The vimentin score was calculated by multiplying both parameters (range, 0-300).

EGFR mutation analysis. Exons 18 through 21 of *EGFR* were polymerase chain reaction (PCR)-amplified using intron-based primers as previously described (15, 16). Approximately 200 microdissected FFPE cells were used for each PCR amplification. All PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and DNA extraction and sequenced in both directions, as previously reported (15, 16).

Statistical analysis. The data were summarized using standard descriptive statistics and frequency tabulation. Associations between categorical variables were assessed via cross-tabulation, the chi-square test, and Fisher's exact test. The Kruskal-Wallis and Wilcoxon rank-sum tests were performed to determine the differences in continuous variables by clinicopathologic feature. Survival curves were estimated using the Kaplan-Meier method. Univariate and multivariate Cox proportional hazard models were used to assess the effect of covariates on overall survival and recurrence-free survival. All computations were carried out using SAS (Cary, NC) and S-plus 2000 (Cambridge, MA) software. *P*-values smaller than 0.05 were considered as statistical significant.

RESULTS

Immunohistochemical expression of EMT markers in lung cancer specimens. Overall, adenocarcinoma and squamous cell carcinoma had high levels of the EMT phenotype on the basis of protein expression of the five markers examined (Figure 1 and Table 2). Both histologic types had significantly reduced levels of E-cadherin membrane expression, as measured by membrane class scores, compared with corresponding adjacent normal and reactive respiratory epithelium (Figure 2). In contrast, lung tumors had significantly higher cytoplasm scores of N-cadherin and MMP-9 and membrane class scores of N-cadherin and integrin-αvβ6 compared with adjacent respiratory epithelium (Figure 2). Vimentin cytoplasm scores were significantly higher in adenocarcinoma tumors than in normal and reactive epithelium but not in squamous cell carcinomas (Figure 2).

NSCLC histologic types had varied patterns of immunohistochemical EMT marker expression (Table 2). E-cadherin membrane expression was low (classes 0 and 1) in 63% of tumors, and this low level of expression was significantly more common in squamous cell carcinomas (73%) than in adenocarcinomas (57%) (*P*=0.005). High levels (classes 2 and 3) of membrane expression of N-cadherin and integrin-αvβ6 were detected in 68% and 56% of NSCLCs, with no differences by histologic type. However, the N-cadherin cytoplasm score was significantly (*P*=0.002) higher in squamous cell carcinoma. No significant differences in the level of cytoplasm expression of vimentin and MMP-9 were detected between adenocarcinoma and squamous cell carcinoma.

We compared membrane class frequencies and scores and cytoplasm scores with patients' demographic and clinicopathologic characteristics, including age, sex, smoking history, and TNM pathologic stage. Interestingly, N-cadherin membrane class and cytoplasm scores were significantly higher (*P*=0.004 and *P*=0.0004, respectively) in ever-smokers than in never-smokers. No significant association was found between any marker expression and pathologic disease stage. We also determined the prognostic (recurrence-free and overall survival) effect of EMT marker expression; the only marker that was associated with outcome was integrin-$\alpha v\beta 6$ membrane expression. As stated above, we previously found (13) that membrane integrin-$\alpha v\beta 6$ is frequently overexpressed in NSCLC tumor specimens using the same TMA set reported here and that this overexpression significantly affected overall survival (data not shown).

To determine whether there was an association between EMT phenotype and EGFR protein expression in NSCLC, we compared the immunohistochemical expression of our five EMT markers with EGFR and pEGFR (Tyr-1086) expression in 225 tumors. No association was detected between the immunohistochemical expression of EMT and the EGFR proteins evaluated. Likewise, tumors' *EGFR* mutation status was not associated with EMT marker expression in a subset of 135 adenocarcinomas, including 32 (24%) mutant cases.

EMT marker immunohistochemical expression in the sequential pathogenesis of squamous cell carcinoma. We determined the expression of five EMT markers in 192 epithelial specimens containing histologically normal, hyperplastic, squamous metaplastic, or squamous dysplastic bronchial epithelia adjacent to squamous cell carcinomas obtained from 89 patients. We found an increased level of the EMT phenotype, corresponding to increased histopathologic severity, in squamous preneoplasias (dysplasias and carcinoma *in situ*) compared with corresponding primary tumors. The membrane expression of E-cadherin was significantly reduced in tumors compared with preneoplastic lesions (Figure 3), but the membrane expression of N-cadherin and integrin-$\alpha v\beta 6$ (Figure 3) and cytoplasm expression of N-cadherin and vimentin (data not shown) were significantly increased. Although squamous preneoplasias had significantly higher expression than did normal and reactive epithelia of membrane integrin-$\alpha v\beta 6$ (Figure 3) and cytoplasm MMP-9 (data not shown), we found that preneoplasias had significantly lower expression of N-cadherin and vimentin than did normal and reactive epithelia (data not shown). Of interest, integrin-αvβ6 membrane expression showed a clear stepwise increase, according to the severity of histopathologic lesions. A high level of expression (classes 2 and 3) was detected in 14 of 73 (19%) normal and reactive epithelia, 18 of 44 (41%) squamous preneoplasias, and 59 of 106 (56%) primary squamous cell carcinomas.

Immunohistochemical expression of EMT markers in brain metastases compared with primary tumors. We studied the expression of EMT markers in 46 NSCLC brain metastases, including 35 adenocarcinomas and 11 squamous cell carcinomas. Interestingly, we found that metastasis sites had a lower level of the EMT phenotype than did primary tumors, which has been suggested to correspond to MET (Table 3) (6). Squamous cell carcinoma brain metastases had a significantly higher E-cadherin membrane class score than did primary tumors. The metastasis sites of both NSCLC histologic types had significantly lower integrin-αvβ6 membrane and N-cadherin membrane and cytoplasm expression scores than did primary tumors (Table 3).

Discussion

To our knowledge, this is the first study to fully characterize the expression of multiple EMT markers in NSCLC primary tumors and metastases and the pathogenesis of squamous cell carcinoma. The use of TMA to assess tumor and epithelial tissue specimens with annotated clinical information, coupled with image analysis of immunohistochemical expression, allowed us to characterize the expression of five EMT markers in a large set of tumor and epithelial specimens comprising the entire spectrum of lung cancer pathogenesis and progression.

First, we demonstrated that adenocarcinoma and squamous cell carcinoma had high levels of the EMT phenotype compared with normal bronchial epithelium. This phenotype is characterized by reduced expression of E-cadherin and overexpression of N-cadherin, integrin-αvβ6, MMP-9, and vimentin. The differential expression of the EMT phenotype between normal and tumor tissues demonstrates that EMT is associated with NSCLC tumor development. The reduced expression of E-cadherin (10, 17, 18) and overexpression of N-cadherin (19), integrin-αvβ6 (13), and MMP-9 (20-22) has been previously reported in NSCLC tissue specimens. Vimentin is considered a hallmark marker for the mesenchymal differentiation of cells, and its expression in NSCLC cell lines has been previously associated with other EMT-related changes (9). However, to our knowledge, our study is the first to report the expression of multiple EMT markers in a large set of NSCLCs using image analysis assessment of immunostaining.

Tumors are defined by their invasive and metastatic potential. This potential may be achieved by tumor cells via EMT. As with the developmental EMT process, multiple EMT-related histologic, molecular, and transcriptional changes occur during carcinoma progression (23). This process is characterized by morphologic modifications, including epithelial polarized cells acquiring a motile non-polarized appearance, and molecular changes, including altered expression of growth and transcription factors and modified expression of cytoskeletal and adhesion molecules (6). The EMT process has been described extensively for several types of cancer, particularly colon, breast, and prostate cancer (8, 24-27). In lung cancer, the loss or switch of adhesion molecules (10, 28) and the expression of classically mesenchymal proteins have been associated with the induction of EMT and carcinoma aggressiveness (9). In addition, the expression of several growth factors and their effectors (29), transcription factors (30, 31), and proteases (32, 33) have been associated with EMT.

Loss of E-cadherin membrane expression and the subsequent reduction of cells’ ability to form stable cell-cell contacts is a hallmarks of EMT (6). In NSCLC, E-cadherin downregulation occurs mostly via epigenetic mechanisms, including gene promoter hypermethylation (34). N-cadherin is mainly expressed in nervous system and mesenchymal normal cells. In epithelial tumors, increased N-cadherin immunohistochemical expression (often *de novo* expression) is accompanied by E-cadherin downregulation (35). In NSCLC, increased N-cadherin membrane and cytoplasm expression has been previously described in approximately one-third of tumors (19). Integrin-αvβ6, an integrin transmembrane glycoprotein, is integrated across the plasma membrane and provides a link between the extracellular matrix and cytoskeletal molecules (36). Integrin-αvβ6 overexpression has been found in other epithelial tumors, including colorectal (24), ovarian (37), and head and neck carcinomas (38). Of interest, integrin-αvβ6 was recently identified by biopanning peptides in NSCLC cell lines (13), and we have reported immunohistochemistry TMA data that validate this method (13). MMPs are part of the proteolytic cascade that degrades the extracellular matrix and allows the migration of tumor and endothelial cells (39). MMP-9 is a gelatinase capable of forming gaps in the basement membrane to facilitate invasion and metastasis (39). High levels of MMP-9 immunohistochemical expression have been reported in nearly 60% of lung tumors (20, 22, 40).

Our findings indicate that the EMT phenotype is homogeneously present in NSCLC; this phenotype is associated with relatively few clinicopathologic features, including tumor histologic characteristics and patients' smoking status. Squamous cell carcinoma had reduced E-cadherin expression and increased N-cadherin cytoplasmic expression compared with adenocarcinoma histology. To our knowledge, only reduced E-cadherin expression in lung squamous cell carcinoma histology has been previously reported (41, 42). Interestingly, in our study, the N-cadherin membrane and cytoplasm scores were significantly higher in ever-smokers than in never-smokers. The biologic reason for this phenomenon is unknown. Several studies have consistently found an association between loss of E-cadherin expression and poor prognosis in

NSCLC patients treated with surgery with curative intent (10, 17, 28, 42). To our knowledge, only one study showed that N-cadherin overexpression was associated with poor outcome in NSCLC patients (in a univariate analysis) (19). We were not able to replicate those results, which could be due to our use of different immunohistochemical methods, TMAs instead of whole tissue sections, or image analysis instead of microscope observation for assessment of immunostaining. MMP-9 immunohistochemical expression has been found to be associated with prognosis in some studies (20, 40, 43) but not in others (22, 43). We had already reported that integrin-αvβ6 immunohistochemical overexpression was associated with poor overall survival in NSCLC patients (13).

Although EMT has been associated with activation of the EGFR pathway (9, 21, 40, 44, 45), we found no relationship between the expression of EMT markers and EGFR immunohistochemical expression or the *EGFR* mutation status of tumors. Of interest, an EMT gene expression signature was identified in NSCLC cell lines that are sensitive to *in vitro* exposure to the EGFR TKI erlotinib (9). Moreover, the re-expression of E-cadherin has been shown to increase the sensitivity of NSCLC cell lines to EGFR TKIs (46). Although a few studies have shown that E-cadherin overexpression of NSCLC tumors determines sensitivity to EGFR TKI (9), no association was been found between *EGFR* mutation and E-cadherin expression in one study (47).

Lung cancer is believed to develop from a series of preneoplastic lesions in the respiratory mucosa, and these abnormalities are frequently extensive and multifocal throughout the respiratory epithelium, indicating a field effect or field cancerization phenomenon (5). These histopathologic changes have been well defined for squamous cell carcinoma, but they are poorly defined for adenocarcinoma (5). Mucosal changes in the large airways that may precede invasive squamous cell carcinoma include squamous dysplasia and carcinoma *in situ* in the central bronchial airway (5). We found increased EMT phenotype levels (especially reduced E-cadherin and increased N-cadherin and integrin-αvβ6 expression), and these levels corresponded with increased histopathologic severity. These findings indicate that EMT occurs early in the pathogenesis of lung cancer. To our knowledge, only reduced E-cadherin expression has been previously described by us (48) and others (49) to commence at the squamous dysplasia stage. Of the five markers studied, membrane integrin-αvβ6 had the clearest stepwise increase in expression, which corresponded to histopathologic severity (from normal and reactive bronchial epithelia to squamous dysplasia and invasive squamous cell carcinoma). Thus, integrin-αvβ6 may be a novel marker for monitoring the progression of centrally located preneoplastic lesions in smokers.

Because of their overexpression in malignant cells and their role in cell survival, motility, and invasion, several EMT-related proteins, including N-cadherin and integrins, have been considered as therapeutic targets in cancer (36, 50). However, the expression of these proteins in the spectrum of advanced NSCLC, including metastasis sites, has not been previously studied. In addition, to our knowledge, this is the first study to assess the expression of EMT markers in lung cancer tissue specimens obtained from metastasis sites. It has been suggested that the progression of epithelial tumors involves spatial and temporal occurrences of EMT, whereby tumor cells acquire a more invasive and metastatic phenotype (6, 8). The disseminated mesenchymal tumor cells must then undergo MET at the site of metastasis because metastasis recapitulates the pathologic characteristics of the cells' corresponding primary tumors. Interestingly, we found lower EMT phenotype levels in lung cancer brain metastases than in primary tumors, confirming that cellular plasticity, which allows cells to undergo EMT and MET in the appropriate microenvironments, is important in metastasis. Our finding of different levels of the EMT phenotype in NSCLC brain metastases than in primary tumors emphasize the importance of characterizing and understanding the molecular events in lung cancer metastasis.

In summary, to our knowledge, this study is the first to comprehensively evaluate the expression of five EMT-related proteins throughout the early development and progression of NSCLC. We demonstrated that the EMT phenotype is frequently expressed in primary squamous cell carcinoma and adenocarcinoma of the lung; EMT is an early phenomenon in the pathogenesis of squamous cell carcinoma, and NSCLC brain metastases show characteristics of MET. Our findings suggest that EMT is a useful target for lung cancer chemoprevention and therapy.

References

1. Jemal A, Siegel R, Ward E, *et al.* Cancer statistics, 2008. *CA Cancer J Clin* 2008; 58(2):71-96.

2. Minna JD, Gazdar A. Focus on lung cancer. *Cancer Cell* 2002; 1:49-52.

3. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC (eds). Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heartedn, Vol., International Agency for Research on Cancer (IARC): Lyon, 2004, pp 9-124.

4. van Zandwijk N. Neoadjuvant strategies for non-small cell lung cancer. *Lung Cancer* 2001; 34 Suppl 2:S145-50.

5. Wistuba, II. Genetics of preneoplasia: lessons from lung cancer. *Curr Mol Med* 2007; 7(1):3-14.

6. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* 2002; 2(6):442-54.

7. Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* 2006; 7(2):131-42.

8. Hugo H, Ackland ML, Blick T, *et al.* Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *J Cell Physiol* 2007; 213(2):374-83.

9. Yauch RL, Januario T, Eberhard DA, *et al.* Epithelial versus mesenchymal phenotype determines in vitro sensitivity and predicts clinical activity of erlotinib in lung cancer patients. *Clin Cancer Res* 2005; 11(24 Pt 1):8686-98.

10. Bremnes RM, Veve R, Hirsch FR, Franklin WA. The E-cadherin cell-cell adhesion complex and lung cancer invasion, metastasis, and prognosis. *Lung Cancer* 2002; 36(2):115-24.

11. Thomson S, Buck E, Petti F, *et al.* Epithelial to mesenchymal transition is a determinant of sensitivity of non-small-cell lung carcinoma cell lines and xenografts to epidermal growth factor receptor inhibition. *Cancer Res* 2005; 65(20):9455-62.

12. Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest* 1997; 111(6):1710-7.

13. Elayadi AN, Samli KN, Prudkin L, *et al.* A peptide selected by biopanning identifies the integrin alphavbeta6 as a prognostic biomarker for nonsmall cell lung cancer. *Cancer Res* 2007; 67(12):5889-95.

14. Wolff AC, Hammond ME, Schwartz JN, *et al.* American Society of Clinical

Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007; 25(1):118-45.

15. Shigematsu H, Lin L, Takahashi T, *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005; 97(5):339-46.

16. Tang X, Shigematsu H, Bekele BN, *et al.* EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res* 2005; 65(17):7568-72.

17. Kase S, Sugio K, Yamazaki K, *et al.* Expression of E-cadherin and beta-catenin in human non-small cell lung cancer and the clinical significance. *Clin Cancer Res* 2000; 6(12):4789-96.

18. Deeb G, Wang J, Ramnath N, *et al.* Altered E-cadherin and epidermal growth factor receptor expressions are associated with patient survival in lung cancer: a study utilizing high-density tissue microarray and immunohistochemistry. *Mod Pathol* 2004; 17(4):430-9.

19. Nakashima T, Huang C, Liu D, *et al.* Neural-cadherin expression associated with angiogenesis in non-small-cell lung cancer patients. *Br J Cancer* 2003; 88(11):1727-33.

20. Cox G, Jones JL, O'Byrne KJ. Matrix metalloproteinase 9 and the epidermal growth factor signal pathway in operable non-small cell lung cancer. *Clin Cancer Res* 2000; 6(6):2349-55.

21. Swinson DE, Cox G, O'Byrne KJ. Coexpression of epidermal growth factor receptor with related factors is associated with a poor prognosis in non-small-cell lung cancer. *Br J Cancer* 2004; 91(7):1301-7.

22. Leinonen T, Pirinen R, Bohm J, *et al.* Expression of matrix metalloproteinases 7 and 9 in non-small cell lung cancer. Relation to clinicopathological factors, beta-catenin and prognosis. *Lung Cancer* 2006; 51(3):313-21.

23. Thompson EW, Newgreen DF, Tarin D. Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition? *Cancer Res* 2005; 65(14):5991-5; discussion 5.

24. Bates RC. The alphaVbeta6 integrin as a novel molecular target for colorectal cancer. *Future Oncol* 2005; 1(6):821-8.

25. Peinado H, Portillo F, Cano A. Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* 2004; 48(5-6):365-75.

26. Tomita K, van Bokhoven A, van Leenders GJ, *et al.* Cadherin switching in human prostate cancer progression. *Cancer Res* 2000; 60(13):3650-4.

27. Gilles C, Polette M, Piette J, *et al.* Vimentin expression in cervical carcinomas: association with invasive and migratory potential. *J Pathol* 1996; 180(2):175-80.

28. Charalabopoulos K, Gogali A, Kostoula OK, Constantopoulos SH. Cadherin superfamily of adhesion molecules in primary lung cancer. *Exp Oncol* 2004; 26(4):256-60.

29. Shintani Y, Maeda M, Chaika N, Johnson KR, Wheelock MJ. Collagen I Promotes EMT in Lung Cancer Cells via TGF-beta3 Signaling. *Am J Respir Cell Mol Biol* 2007.

30. Shih JY, Tsai MF, Chang TH, *et al.* Transcription repressor slug promotes carcinoma invasion and predicts outcome of patients with lung adenocarcinoma. *Clin Cancer Res* 2005; 11(22):8070-8.

31. Dohadwala M, Yang SC, Luo J, *et al.* Cyclooxygenase-2-dependent regulation of E-cadherin: prostaglandin E(2) induces transcriptional repressors ZEB1 and snail in non-small cell lung cancer. *Cancer Res* 2006; 66(10):5338-45.

32. Schutz A, Schneidenbach D, Aust G, *et al.* Differential expression and activity status of MMP-1, MMP-2 and MMP-9 in tumor and stromal cells of squamous cell carcinomas of the lung. *Tumour Biol* 2002; 23(3):179-84.

33. Illman SA, Lehti K, Keski-Oja J, Lohi J. Epilysin (MMP-28) induces TGF-beta mediated epithelial to mesenchymal transition in lung carcinoma cells. *J Cell Sci* 2006; 119(Pt 18):3856-65.

34. Zochbauer-Muller S, Fong KM, Virmani AK, *et al.* Aberrant promoter methylation of multiple genes in non-small cell lung cancers. *Cancer Res* 2001; 61(1):249-55.

35. Cavallaro U, Christofori G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer* 2004; 4(2):118-32.

36. Eble JA, Haier J. Integrins in cancer treatment. *Curr Cancer Drug Targets* 2006; 6(2):89-105.

37. Ahmed N, Thompson EW, Quinn MA. Epithelial-mesenchymal interconversions in normal ovarian surface epithelium and ovarian carcinomas: an exception to the norm. *J Cell Physiol* 2007; 213(3):581-8.

38. Regezi JA, Ramos DM, Pytela R, Dekker NP, Jordan RC. Tenascin and beta 6 integrin are overexpressed in floor of mouth in situ carcinomas and invasive squamous cell carcinomas. *Oral Oncol* 2002; 38(4):332-6.

39. Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* 2003; 3(5):362-74.

40. Ferrigan L, Wallace WA. Predicting non-small cell lung cancer expression of epidermal

growth factor receptor and matrix metalloproteinase 9 from immunohistochemical staining of diagnostic biopsy samples. *Eur J Cancer* 2004; 40(10):1589-92.

41. Liu D, Huang C, Kameyama K, *et al.* E-cadherin expression associated with differentiation and prognosis in patients with non-small cell lung cancer. *Ann Thorac Surg* 2001; 71(3):949-54; discussion 54-5.

42. Choi YS, Shim YM, Kim SH, *et al.* Prognostic significance of E-cadherin and beta-catenin in resected stage I non-small cell lung cancer. *Eur J Cardiothorac Surg* 2003; 24(3):441-9.

43. Iniesta P, Moran A, De Juan C, *et al.* Biological and clinical significance of MMP-2, MMP-9, TIMP-1 and TIMP-2 in non-small cell lung cancer. *Oncol Rep* 2007; 17(1):217-23.

44. Fedor-Chaiken M, Hein PW, Stewart JC, Brackenbury R, Kinch MS. E-cadherin binding modulates EGF receptor activation. *Cell Commun Adhes* 2003; 10(2):105-18.

45. Andl CD, Rustgi AK. No one-way street: cross-talk between e-cadherin and receptor tyrosine kinase (RTK) signaling: a mechanism to regulate RTK activity. *Cancer Biol Ther* 2005; 4(1):28-31.

46. Witta SE, Gemmill RM, Hirsch FR, *et al.* Restoring E-cadherin expression increases sensitivity to epidermal growth factor receptor inhibitors in lung cancer cell lines. *Cancer Res* 2006; 66(2):944-50.

47. Miyanaga A, Gemma A, Ando M, *et al.* E-cadherin expression and epidermal growth factor receptor mutation status predict outcome in non-small cell lung cancer patients treated with gefitinib. *Oncol Rep* 2008; 19(2):377-83.

48. Shen J, Behrens C, Wistuba, II, *et al.* Identification and validation of differences in protein levels in normal, premalignant, and malignant lung cells and tissues using high-throughput Western Array and immunohistochemistry. *Cancer Res* 2006; 66(23):11194-206.

49. Kato Y, Hirano T, Yoshida K, *et al.* Frequent loss of E-cadherin and/or catenins in intrabronchial lesions during carcinogenesis of the bronchial epithelium. *Lung Cancer* 2005; 48(3):323-30.

50. Mariotti A, Perotti A, Sessa C, Ruegg C. N-cadherin as a therapeutic target in cancer. *Expert Opin Investig Drugs* 2007; 16(4):451-65.

Figure Legends

Figure 1. Representative microphotographs of immunohistochemical expression of EMT markers in tissue specimens of primary NSCLC adenocarcinoma and squamous cell carcinoma. Tumor cells show membrane immunostaining for E-cadherin and integrin-αvβ6, membrane and cytoplasm staining for N-cadherin, and cytoplasm staining for vimentin and MMP-9.

Figure 2. Scores of membrane and cytoplasm immunohistochemical expression of EMT markers in normal and reactive epithelia compared with those in primary tumors. The number of samples is indicated for each histologic group and marker. *P* values comparing normal epithelial and tumor histologic types are shown for all comparisons.

Figure 3. Scores for membrane immunohistochemical expression of E-cadherin, N-cadherin, and integrin-αvβ6 in bronchial respiratory epithelial lesions related to the pathogenesis of squamous cell carcinoma of the lung: normal and reactive epithelium, squamous preneoplasia (dysplasias and carcinoma *in situ*), and primary squamous cell carcinoma. The number of samples is indicated for each histologic group and marker. Significant *P* values for comparisons between each group are shown. For each marker, representative microphotographs of immunohistochemical expression in tissue specimens of bronchial epithelium with normal, squamous dysplasia, and carcinoma *in situ* are shown. For E-cadherin, decreased membrane immunostaining was found with increased histologic severity. The opposite phenomenon is shown for cytoplasm N-cadherin and membrane integrin-αvβ6 expression.

Table 1. Demographic and clinicopathologic data by histologic tumor type

Characteristic	**Total (N=325)**	**Adenocarcinoma (N=209)**	**Squamous cell carcinoma (N=116)**
Sex			
Female	173	128	45
Male	152	81	71
Stage			
I	205	139	66
II	62	28	34
III	49	35	14
IV	9	7	2
Smoking history[1]			
Yes	263	152	111
No	61	57	4
Smoking status[1]			
Never	61	57	4
Former	161	92	69
Current	102	60	42

[1] Smoking history and status was not available for one squamous cell carcinoma patient.

Table 2. Immunohistochemical expression of EMT markers by NSCLC histologic type

Marker	Adenocarcinoma	Squamous cell carcinoma	P value	Total
E-cadherin				
Membrane class				
N	191	110		301
0 or 1 (%)	108 (57)	80 (73)	0.005	188 (62)
2 or 3 (%)	83 (43)	30 (27)		113 (38)
N-cadherin				
Membrane class				
N	193	106	0.73	299
0 or 1 (%)	60 (31)	35 (33)		95 (32)
2 or 3 (%)	133 (69)	71 (67)		204 (68)
Mean cytoplasm score (SD)	46.4 (4.2)	47.8 (2.4)	0.002	46.9 (3.7)
Integrin-αvβ6				
Membrane class				
N	177	106	0.89	283
0 or 1 (%)	77 (44)	47 (44)		124 (44)
2 or 3 (%)	100 (56)	59 (56)		159 (56)
Vimentin				
Cytoplasm score				
N	172	90		262
Mean (SD)	62.1 (92.6)	42.2 (72.4)	0.15	55.3 (86.6)
MMP-9				
Cytoplasm score				
N	179	96		275
Mean (SD)	44.8 (9.9)	44.9 (8.4)	0.95	44.8 (9.4)

Table 3. Immunohistochemical expression of EMT markers in primary tumors and brain metastases by histologic type

	Adenocarcinoma			Squamous cell carcinoma		
Marker	**Primary tumor**	**Metastasis**	**P value**	**Primary tumor**	**Metastasis**	**P value**
E-cadherin						
Membrane class						
N	191	35		110	11	
Mean (SD)	1.5 (0.6)	1.6 (0.7)	0.16	1.3 (0.5)	1.6 (0.8)	0.02
N-cadherin						
Membrane class						
N	193	37		106	11	
Mean (SD)	1.7 (0.6)	1.2 (0.9)	<0.0002	1.7 (0.5)	1 (0.4)	<0.0002
Cytoplasm score						
Mean (SD)	46.4 (4.2)	20.5 (20.1)	<0.0002	47.8 (2.4)	28.6 (20.7)	<0.0002
Integrin-αvβ6						
Membrane class						
N	177	32		106	11	
Mean (SD)	1.8 (0.9)	1.2 (0.9)	<0.0002	1.8 (0.9)	1.5 (0.9)	0.04
Vimentin						
Cytoplasm score						
N	172	32		90	10	
Mean (SD)	62.1 (92.6)	61.3 (101.3)	0.9	42.2 (72.4)	15 (30.6)	0.04
MMP-9						
Cytoplasm score						
N	179	31		96	8	
Mean (SD)	44.8 (9.9)	47.1 (2.9)	0.07	44.9 (8.4)	47.4 (1.6)	0.12

Figure 1

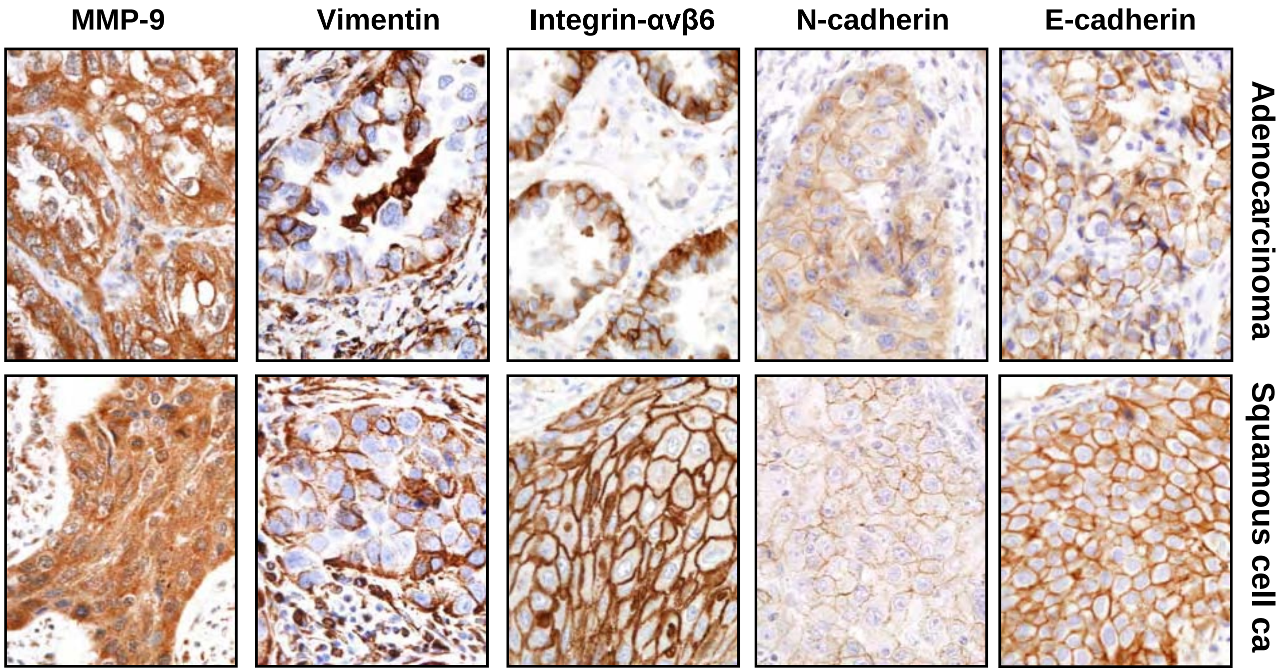

MMP-9
Vimentin
Integrin-αvβ6
N-cadherin
E-cadherin
Adenocarcinoma
Squamous cell ca

Figure 2

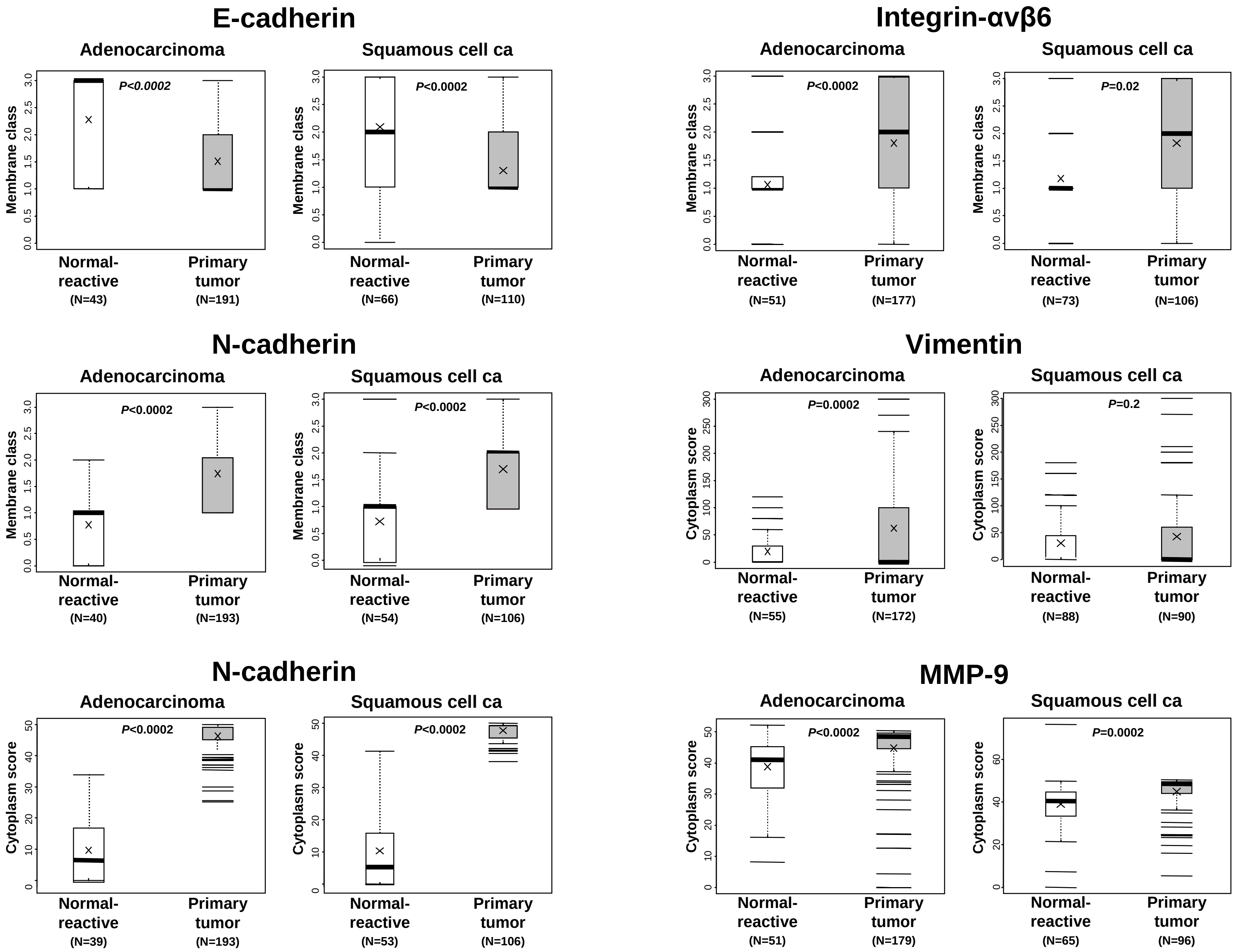
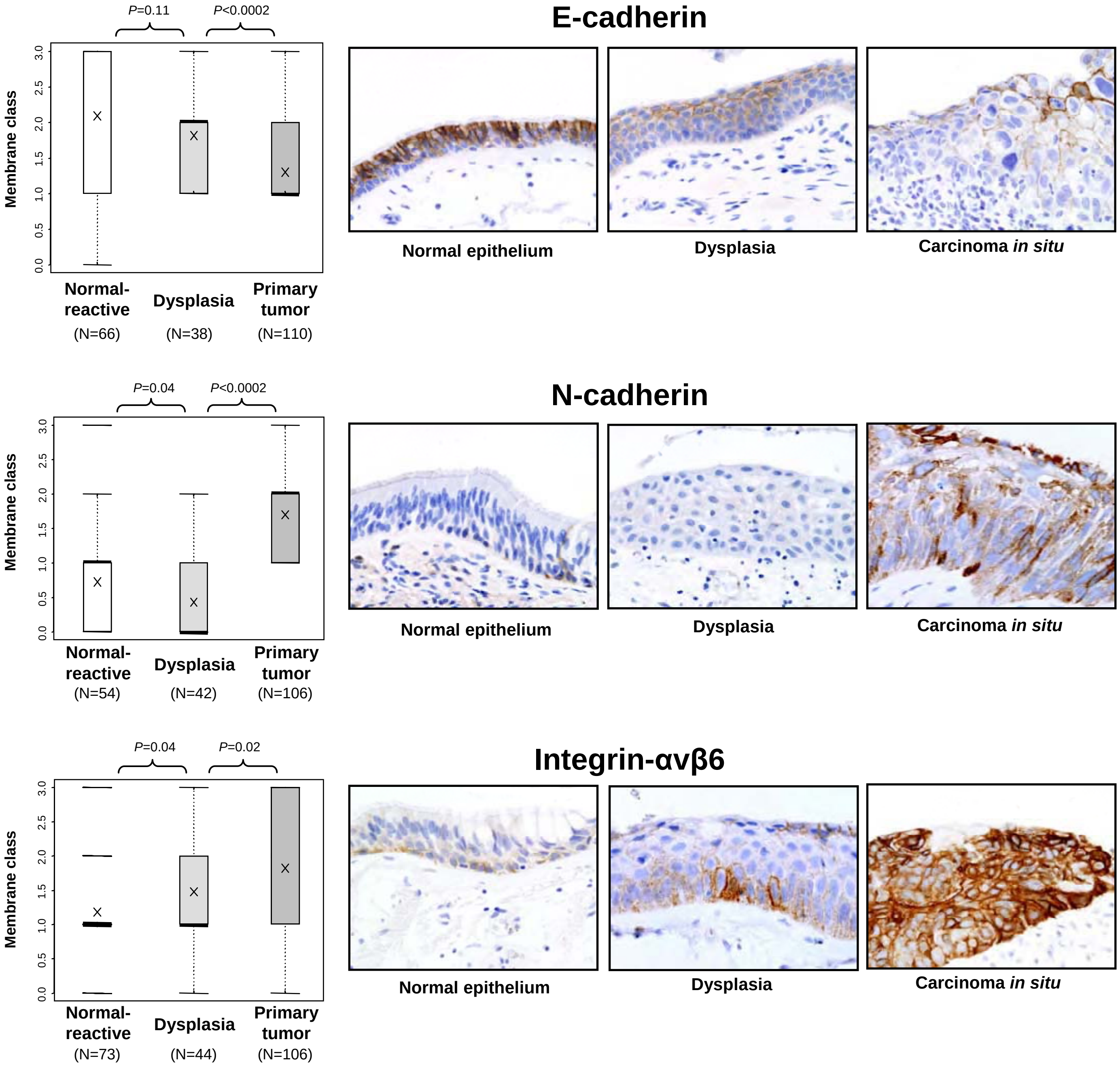

E-cadherin
Adenocarcinoma
Squamous cell ca
Membrane class
P<0.0002
P<0.0002
Normal-reactive (N=43)
Primary tumor (N=191)
Normal-reactive (N=66)
Primary tumor (N=110)
Integrin-αvβ6
Adenocarcinoma
Squamous cell ca
Membrane class
P<0.0002
P=0.02
Normal-reactive (N=51)
Primary tumor (N=177)
Normal-reactive (N=73)
Primary tumor (N=106)
N-cadherin
Adenocarcinoma
Squamous cell ca
Membrane class
P<0.0002
P<0.0002
Normal-reactive (N=40)
Primary tumor (N=193)
Normal-reactive (N=54)
Primary tumor (N=106)
Vimentin
Adenocarcinoma
Squamous cell ca
Cytoplasm score
P=0.0002
P=0.2
Normal-reactive (N=55)
Primary tumor (N=172)
Normal-reactive (N=88)
Primary tumor (N=90)
N-cadherin
Adenocarcinoma
Squamous cell ca
Cytoplasm score
P<0.0002
P<0.0002
Normal-reactive (N=39)
Primary tumor (N=193)
Normal-reactive (N=53)
Primary tumor (N=106)
MMP-9
Adenocarcinoma
Squamous cell ca
Cytoplasm score
P<0.0002
P=0.0002
Normal-reactive (N=51)
Primary tumor (N=179)
Normal-reactive (N=65)
Primary tumor (N=96)

Figure 3


E-cadherin
P=0.11
P<0.0002
Membrane class
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Normal-reactive
(N=66)
Dysplasia
(N=38)
Primary tumor
(N=110)
Normal epithelium
Dysplasia
Carcinoma in situ
N-cadherin
P=0.04
P<0.0002
Membrane class
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Normal-reactive
(N=54)
Dysplasia
(N=42)
Primary tumor
(N=106)
Normal epithelium
Dysplasia
Carcinoma in situ
Integrin-αvβ6
P=0.04
P=0.02
Membrane class
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Normal-reactive
(N=73)
Dysplasia
(N=44)
Primary tumor
(N=106)
Normal epithelium
Dysplasia
Carcinoma in situ

Research Article

Cyclic AMP Response Element-Binding Protein Overexpression: A Feature Associated with Negative Prognosis in Never Smokers with Non–Small Cell Lung Cancer

Hye-Sook Seo,[1] Diane D. Liu,[2] B. Nebiyou Bekele,[2] Mi-Kyoung Kim,[1] Katherine Pisters,[1]
Scott M. Lippman,[1] Ignacio I. Wistuba,[1,3] and Ja Seok Koo[1]

Departments of [1]Thoracic/Head and Neck Medical Oncology, [2]Biostatistics, and [3]Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Abstract

Lung cancer is the leading cause of cancer deaths worldwide. Recent advances in targeted therapies hold promise for the development of new treatments for certain subsets of cancer patients by targeting specific signaling molecule. Based on the identification of the transcription factor cyclic AMP response element-binding protein (CREB) as an important regulator of growth of several types of cancers and our recent findings of its importance in normal differentiation of bronchial epithelial cells, we hypothesized that CREB plays an important pathobiologic role in lung carcinogenesis. We conducted this initial study to determine whether the expression and activation status of CREB are altered in non–small cell lung cancer (NSCLC) and of any prognostic importance in NSCLC patients. We found that the expression levels of mRNA and protein of CREB and phosphorylated CREB (p-CREB) were significantly higher in most of the NSCLC cell lines and tumor specimens than in the normal human tracheobronchial epithelial cells and adjacent normal lung tissue, respectively. Analysis of CREB mRNA expression and the *CREB* gene copy number showed that CREB overexpression occurred mainly at the transcriptional level. Immunohistochemical analysis of tissue microarray slides containing sections of NSCLC specimens obtained from 310 patients showed that a decreased survival duration was significantly associated with overexpression of CREB or p-CREB in never smokers but not in current or former smokers with NSCLC. These are the first reported results illustrating the potential of CREB as a molecular target for the prevention and treatment of NSCLC, especially in never smokers. [Cancer Res 2008;68(15):6065–73]

Introduction

Lung cancer is the leading cause of cancer deaths, and its incidence is rising (1). In the United States, 215,020 new cases and 161,840 deaths of lung and bronchial cancer are projected to occur in 2008 (2). Non–small cell lung cancer (NSCLC) accounts for ~80% of all lung cancers and is subdivided into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma (3). Recent advances in targeted therapies directed against the epidermal growth factor receptor (EGFR) and vascular endothelial growth factor receptor pathways showed marked improvement in treatment in a subset of patients with lung cancer (4–6). Thus, identifying new molecular targets for treatment and/or prevention of NSCLC is warranted and urgently needed to improve the control of this deadly form of lung cancer.

Studies have shown that cyclic AMP (cAMP) response element-binding protein (CREB) plays important roles in cell differentiation (7), survival (8, 9), proliferation (10), development (11, 12), cell cycle progression (10), and glucose metabolism (13). CREB is activated by cAMP, growth factors, hormones, retinoids (14), cytokines (15), and prostaglandins (16) via multiple signaling pathways (13, 17), including the cAMP/protein kinase A, phosphatidylinositol 3-kinase (PI3K)/Akt, extracellular signal-regulated kinase (ERK)/p90 ribosomal S6 kinase, and p38/mitogen- and stress-activated protein kinase pathways. Once activated, CREB induces the expression of cAMP response element-containing target genes (13, 18, 19), which play important roles in differentiation (14), cell cycle progression (20), apoptosis suppression (21), proliferation (22), neovascularization (23), inflammation (24), and tumorigenesis (25).

Recently, we showed that CREB plays a physiologic role in mucous differentiation of normal human tracheobronchial epithelial (NHTBE) cells (14, 26). Previous studies found that CREB has a pathobiologic role in the growth of breast cancer, melanoma, and hepatocellular carcinoma cells (7, 27, 28). Recent studies also showed that CREB acts as a proto-oncogene to regulate hematopoiesis and to contribute to the leukemia phenotype (29–31). However, whether CREB expression is altered in human NSCLC tumors and whether CREB/phosphorylated CREB (p-CREB) expression correlates with the survival rate in patients with NSCLC have not been previously shown.

Based on these previous findings, we hypothesized that CREB, which plays an important role in normal differentiation of bronchial epithelial cells, may also have an important pathobiologic role in lung carcinogenesis as a transcriptional regulatory factor. To test this hypothesis, we compared the CREB expression levels and activation statuses and the *CREB* gene copy numbers in 10 NSCLC cell lines, NHTBE cells, 6 frozen human NSCLC tissue specimens, and paired normal lung tissue specimens. We also analyzed CREB and p-CREB expression in 45 paraffin-embedded whole specimens of NSCLC tumor and adjacent normal bronchial or bronchiolar epithelial tissue specimens. Lastly, we assessed the levels of CREB and p-CREB expression in association with clinicopathologic variables and overall survival duration of 310 NSCLC patients with banked NSCLC tissue specimens using tissue microarray (TMA) analysis. Although studies have consistently shown that smoking is an important etiologic factor for lung

---

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Current address for M-K. Kim: Department of Pathology, Chung-Ang University Medical Center, Seoul, South Korea.

Requests for reprints: Ja Seok Koo, Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Unit 432, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-8454; Fax: 713-794-5997; E-mail: jskoo@mdanderson.org.

©2008 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-5376

cancer, about 15% of men and 53% of women with this disease worldwide are never smokers (32). Thus, we also analyzed the effects of CREB and p-CREB expression on overall survival duration in patients with NSCLC according to smoking status.

Materials and Methods

NSCLC tissue specimens and TMA construction. Six frozen tumor tissue specimens (three squamous cell carcinoma and three adenocarcinoma) and six adjacent normal lung tissue specimens surgically resected from patients who underwent lobectomies or pneumonectomies for primary NSCLC were obtained from the tissue bank of The University of Texas M. D. Anderson Cancer Center Specialized Program of Research Excellence in Lung Cancer. All of the tumors were histologically examined and classified using the 2004 WHO International Classification of Lung Tumors (33). In addition, specimens of tumor and adjacent normal lung tissue (including bronchial and bronchiolar epithelia) obtained from 45 patients with surgically resected NSCLC (26 adenocarcinoma and 19 squamous cell carcinoma) were randomly selected for assessment of immunohistochemical expression of CREB and p-CREB in whole histologic sections. After histologic examination of 310 NSCLC specimens [194 adenocarcinoma or bronchioloalveolar carcinoma (BAC) and 116 squamous cell carcinoma] in the tissue bank of The University of Texas M. D. Anderson Cancer Center, tumor TMAs were constructed using three tissue cores per tumor that were 1 mm in diameter to obtain tissue from central, intermediate, and peripheral tumor areas. The M. D. Anderson Cancer Center Institutional Review Board approved the use of the archived clinical tissue specimens.

Smoking history. Patients who had smoked at least 100 cigarettes in their lifetime were defined as ever smokers, and patients who quit smoking at least 12 mo before their lung cancer diagnosis were defined as former smokers. Current smokers were defined as active smokers who had been smoking for at least 6 mo. Subjects were asked whether they had ever smoked any tobacco products (nonfiltered or filtered cigarettes, cigars, or pipes) for at least 6 mo and were classified as never smokers if they had not,

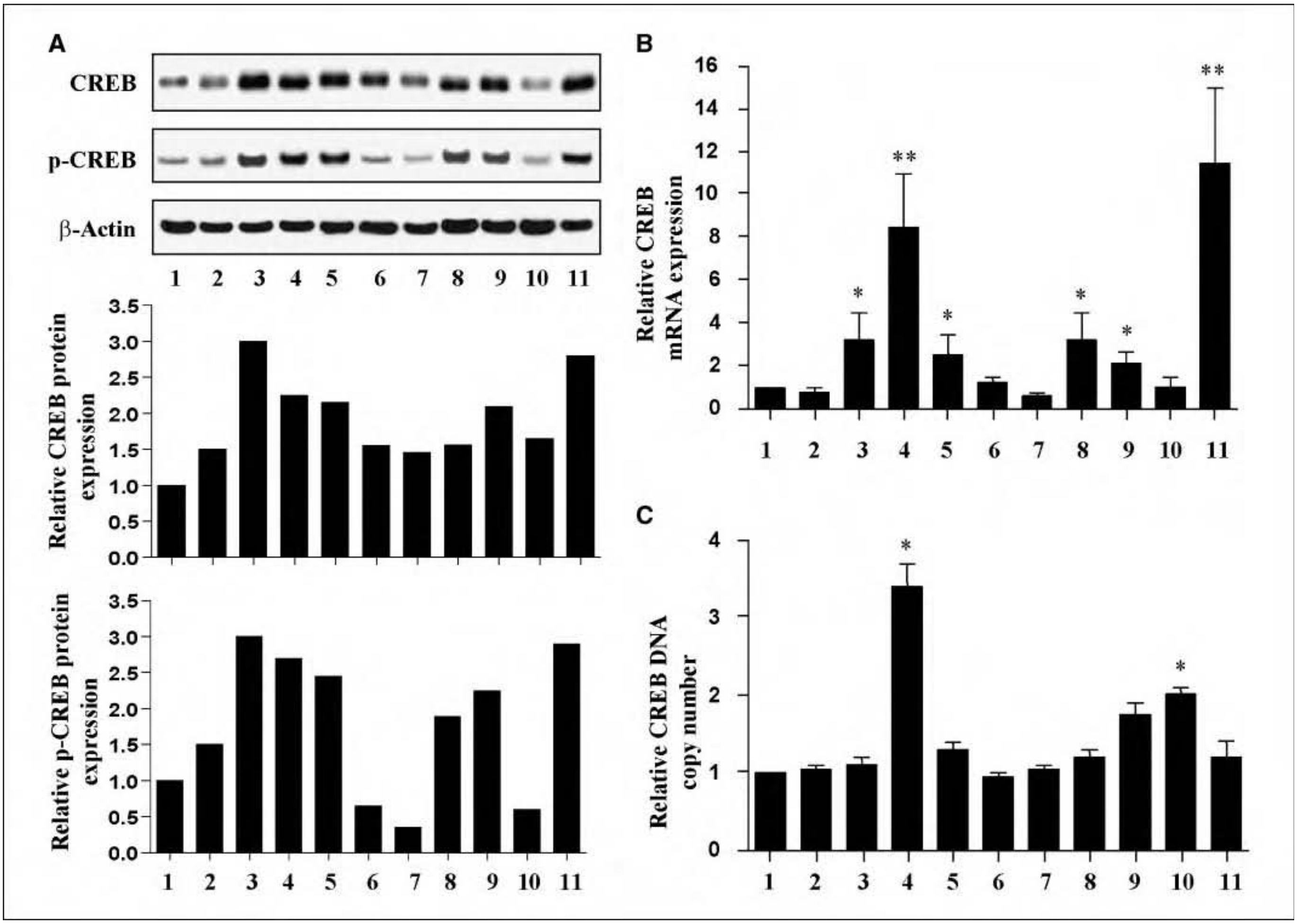


Figure 1. Expression of CREB and p-CREB in NHTBE and NSCLC cells. *A,* Western blot analysis of CREB and p-CREB expression. Whole-cell lysates were prepared from fully differentiated NHTBE cells (*lane 1*) and the 10 NSCLC cell lines (*lane 2*, H226; *lane 3*, H292; *lane 4*, H520; *lane 5*, H2170; *lane 6*, H1563; *lane 7*, H1734; *lane 8*, H1975; *lane 9*, H2228; *lane 10*, A549; *lane 11*, H1703) grown in optimal medium up to confluence. Equal amounts of lysate were subjected to SDS-PAGE and Western blotting. The blots were probed with anti-CREB and anti-p-CREB antibodies. Equal protein loading was confirmed by stripping the blots and reprobing them with an anti-β-actin antibody. The expression levels of CREB and p-CREB proteins for each cell line in relation to the NHTBE cells were quantitated. *B,* qRT-PCR analysis of CREB mRNA expression. The CREB mRNA level in each cell line was analyzed with qRT-PCR. The values shown are the ratios of the CREB mRNA expressed in NSCLC cells to that expressed in NHTBE cells, with CREB mRNA levels normalized against the GAPDH mRNA level. The results shown are from a representative experiment performed twice, each run in triplicate. *Columns,* mean; *bars,* SE. *, $P < 0.05$; **, $P < 0.01$ versus NHTBE cells (*lane 1*) with Student's *t* test. *C, CREB* gene copy number analysis with PCR. The genomic DNA obtained from each cell line was subjected to PCR. The results shown are from a representative experiment performed twice, each run in triplicate. The values are the ratios of the *CREB* DNA copy number in NSCLC cells to that in NHTBE cells, with CREB DNA normalized against the β-actin DNA level. *Columns,* mean; *bars,* SE. *, $P < 0.05$; **, $P < 0.01$ versus NHTBE cells (*lane 1*) with Student's *t* test.

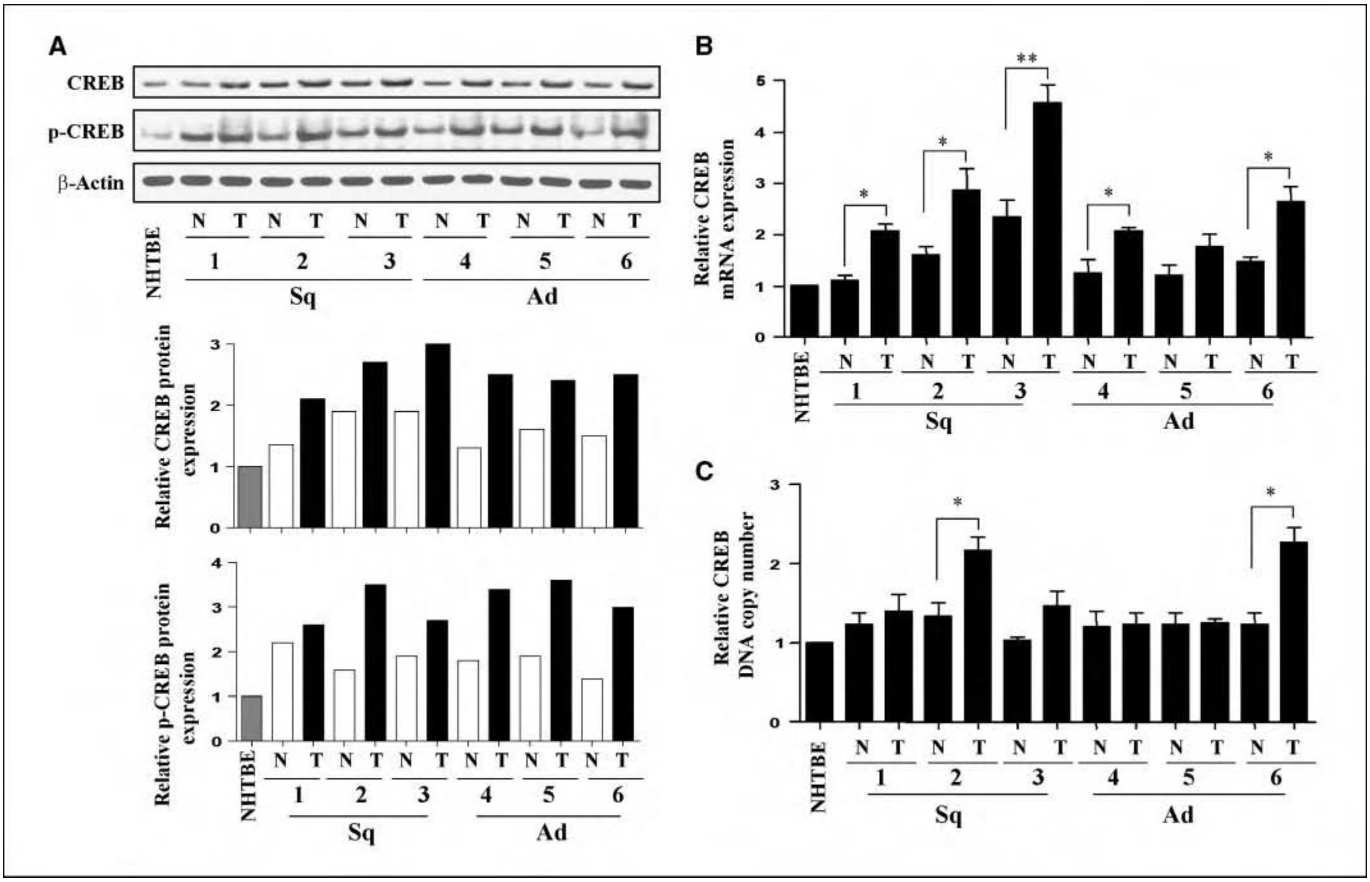


Figure 2. Expression of CREB and p-CREB in frozen human NSCLC specimens and adjacent normal bronchial and bronchiolar epithelial tissue specimens. *A,* Western blot analysis of CREB and p-CREB expression. Soluble proteins obtained from three squamous cell carcinoma (*Sq*) and three adenocarcinoma (*Ad*) tissue specimens (*T*) and paired matching normal tissue specimens (*N*) were subjected to Western blot analysis for the levels of total CREB and p-CREB expression. Equal protein loading was confirmed by stripping the blots and reprobing them with an anti-β-actin antibody. The expression levels of CREB and p-CREB proteins in tissue specimens in relation to that of the NHTBE cells (*hatched columns*) were quantitated. *B,* qRT-PCR analysis of CREB mRNA expression. Total mRNA from the NSCLC tissue specimens and paired normal tissue specimens described in *A* was subjected to qRT-PCR. The values shown are the ratios of the CREB mRNA expressed in tissue specimens to that expressed in NHTBE cells, with CREB mRNA levels normalized against the GAPDH mRNA level. The results are from a representative experiment performed twice, and samples were run in triplicate. *Columns,* mean; *bars,* SE. *, $P < 0.05$; **, $P < 0.01$ versus matched normal tissue with Student's *t* test. *C,* PCR analysis of *CREB* gene copy number. The genomic DNA extracted from the NSCLC tissue specimens and paired normal tissue specimens was subjected to quantitative PCR. The values are the ratios of the *CREB* DNA copy numbers in normal and tumor tissues to the *CREB* DNA copy numbers in NHTBE cells, with *CREB* DNA normalized against the β-actin DNA level. The results are from a representative experiment performed twice, and samples were run in triplicate. *Columns,* mean; *bars,* SE. *, $P < 0.05$; **, $P < 0.01$ versus matched normal tissue with Student's *t* test.

as current smokers if they smoked daily at the date of diagnosis (for patients) or interview (for controls), or as former smokers if they had stopped smoking daily before those dates.

Cell cultures. NHTBE cells obtained from Cambrex were organotypically cultured and maintained as described previously (34–37). Basically, NHTBE cells from passage 2 were seeded at a density of 1×10^5 per insert onto 24-mm, uncoated, semipermeable membranes (Transwell clear; Costar) in a 1:1 mixture of DMEM (Invitrogen Co.) and bronchial epithelial cell basal medium (Cambrex) supplemented with transferrin (10 ng/mL), epinephrine (0.5 μg/mL), insulin (5 μg/mL), triiodothyronine (6.5 ng/mL), hydrocortisone (0.5 μg/mL), EGF (0.5 ng/mL), bovine pituitary extract (1% w/v), bovine serum albumin (1.5 μg/mL), gentamicin (10 μg/mL), and retinoic acid (5×10^{-8} mol/L). The cells were grown submerged in the medium for the first 7 d, after which time an air-liquid interface was created. The cells were then cultured in the air-liquid interface for 3 wk, with the medium changed every 24 h. Fully differentiated 28-d-old cultures developed mucociliary phenotypes similar to that of *in vivo* bronchial epithelium. In addition, 10 NSCLC cell lines (H226, H292, H520, H2170, H1563, H1734, H1975, H2228, A549, and H1703) were obtained from the American Type Culture Collection and maintained in RPMI 1640 containing 10% fetal bovine serum and gentamicin (10 μg/mL).

Western blot analysis. Western blot analysis was performed as described previously (14) to measure the expression of CREB and p-CREB in whole-cell extracts from the NHTBE cells and NSCLC cell lines and from the six archived NSCLC specimens and paired normal lung tissue specimens.

CREB mRNA expression and *CREB* gene copy number. Total RNA and genomic DNA from the NHTBE cells, NSCLC cells, and frozen NSCLC specimens were extracted using the Qiagen RNeasy Mini kit and Blood & Cell Culture DNA Mini kit (Qiagen) and subjected to quantitative reverse transcription-PCR (qRT-PCR) and conventional PCR analysis to determine CREB mRNA expression and *CREB* gene copy number, as described previously (14). The primer sequences used for detection of CREB mRNA in qRT-PCR were as follows: forward, 5′-ACTGTAACGGTGCCAACTCC-3′; reverse, 5′-GAATGGTAGTACCCGGCTGA-3′. The mRNA level of human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) detected with VIC dye (Applied Biosystems) was used as endogenous control. The primer sequences used for determination of the *CREB* gene copy number in conventional PCR were as follows: forward, 5′-AAGAGGAGACTTCAGCACCTG-3′; reverse, 5′-GCAAAACTGAGAAGACTTGGC-3′. For endogenous control, the following β-actin primer sequences were used: forward, 5′-AGGTCATCACCATTGGCAAT-3′; reverse, 5′-AATGAGGGCAGGACTTAGCTT-3′.

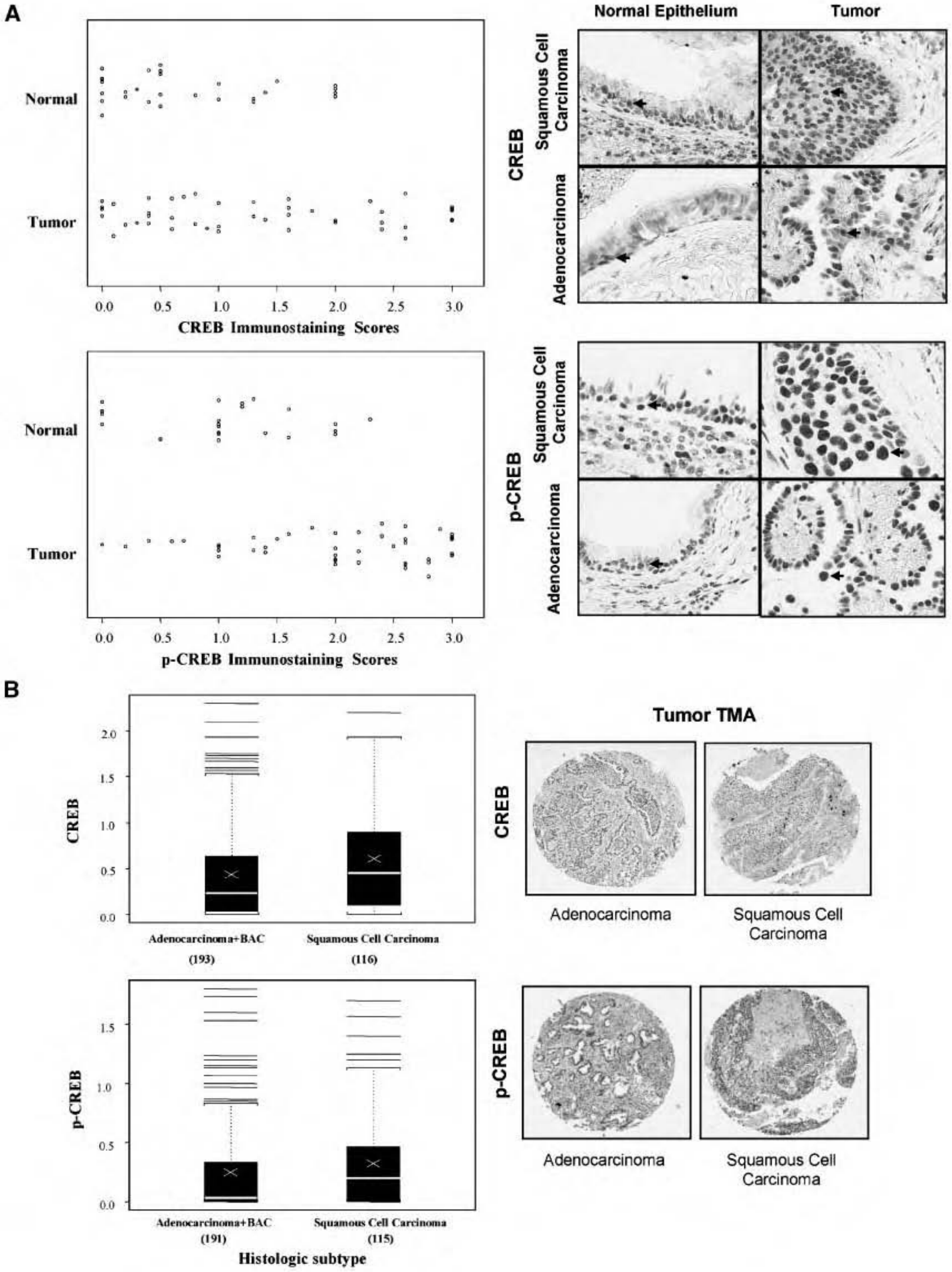

A
Normal Epithelium
Tumor
CREB
Squamous Cell Carcinoma
Adenocarcinoma
Normal
Tumor
0.0
0.5
1.0
1.5
2.0
2.5
3.0
CREB Immunostaining Scores
p-CREB
Squamous Cell Carcinoma
Adenocarcinoma
Normal
Tumor
0.0
0.5
1.0
1.5
2.0
2.5
3.0
p-CREB Immunostaining Scores
B
CREB
2.0
1.5
1.0
0.5
0.0
Adenocarcinoma+BAC
(193)
Squamous Cell Carcinoma
(116)
Tumor TMA
CREB
Adenocarcinoma
Squamous Cell Carcinoma
p-CREB
1.5
1.0
0.5
0.0
Adenocarcinoma+BAC
(191)
Squamous Cell Carcinoma
(115)
Histologic subtype
p-CREB
Adenocarcinoma
Squamous Cell Carcinoma

Figure 3. Immunohistochemical analysis of CREB and p-CREB expression in normal and tumor lung tissues. *A,* fixed NSCLC tissue specimens (26 adenocarcinoma and 19 squamous cell carcinoma) and adjacent normal bronchial and bronchiolar epithelial tissue specimens were subjected to immunohistochemical staining and then scored according to the criteria mentioned in Materials and Methods. *Left,* BLiP plots displaying the distribution of CREB and p-CREB immunostaining scores in normal and tumor tissues. Average scores for CREB are 1.37 in tumor versus 0.73 in normal ($P = 0.013$) and for p-CREB are 1.96 in tumor versus 1.05 in normal ($P = 0.0002$). Representative images (*right*) were captured at a magnification of ×200. *Arrows,* nuclear CREB and p-CREB immunostaining in the basal layer of normal bronchial epithelial and tumor tissue specimens. *B,* CREB and p-CREB immunostaining of TMA specimens from a total of 310 patients with either adenocarcinoma (194) or squamous cell carcinoma (116) are included in the analysis. *Left,* the distributions of the staining scores for the two histologic types of NSCLC are presented in the box plots. The number of samples measured was indicated under each category. X marks and lines inside the quartile boxes are means and medians, respectively (for CREB, 0.43 and 0.23 in adenocarcinoma/BAC versus 0.61 and 0.45 in squamous cell carcinoma; for p-CREB, 0.25 and 0.03 versus 0.32 and 0.2). $P = 0.002$ for CREB; $P = 0.008$ for p-CREB. *Right,* representative images of CREB and p-CREB immunostaining of TMA NSCLC specimens. Magnification, ×40.

Table 1. Estimation of overall survival durations per demographic and clinicopathologic characteristics in NSCLC patients from whom tumor TMA specimens were obtained

Characteristic	Variable estimate ± SE	*P*	Hazards ratio
Univariate Cox proportional hazards regression models			
Age	0.04 ± 0.01	0.0002	1.04
Sex (male vs female)	0.34 ± 0.21	0.1000	1.40
Race (white vs other)	−0.31 ± 0.34	0.3500	0.73
Smoking status			
Former vs never	0.24 ± 0.25	0.3300	1.27
Current vs never	0.03 ± 0.29	0.9300	1.03
Histologic subtype (squamous cell carcinoma vs adenocarcinoma/BAC)	0.65 ± 0.21	0.0020	1.92
CREB expression	0.28 ± 0.17	0.1000	1.33
p-CREB expression	0.66 ± 0.23	0.0040	1.93
Pathologic T classification (T2 + T3 + T4 vs T1)	0.97 ± 0.27	0.0002	2.63
Pathologic N classification (N1 + N2 + N3 vs N0)	0.73 ± 0.23	0.0010	2.08
Pathologic M classification (M1 vs M0)	1.05 ± 0.72	0.1400	2.87
Pathologic TNM stage			
II vs I	0.65 ± 0.26	0.0100	1.92
III + IV vs I	0.86 ± 0.29	0.0030	2.37
Multivariate Cox proportional hazards regression models			
Model A: CREB			
Age	0.04 ± 0.01	0.0002	1.04
CREB expression	0.41 ± 0.18	0.0200	1.51
Pathologic T classification (T2 + T3 + T4 vs T1)	0.84 ± 0.28	0.0030	2.31
Pathologic N classification (N1 + N2 + N3 vs N0)	0.47 ± 0.24	0.0490	1.60
Model B: p-CREB			
Age	0.04 ± 0.01	0.0005	1.04
p-CREB expression	0.59 ± 0.23	0.0100	1.80
Pathologic T classification (T2 + T3 + T4 vs T1)	0.69 ± 0.28	0.0100	2.00
Pathologic N classification (N1 + N2 + N3 vs N0)	0.47 ± 0.24	0.0500	1.59

Immunohistochemical analysis. Immunohistochemical analysis of the NSCLC and adjacent normal bronchial and bronchiolar epithelial tissue specimens and of the TMA NSCLC tissue specimens was performed using anti-CREB and anti-p-CREB antibodies (Upstate) at a dilution of 1:100 according to the manufacturer's instructions. Immunostaining was visualized using the Histostain-Bulk-SP kit and the AEC red substrate kit (Zymed Laboratories). Immunohistochemical staining without a primary antibody was performed as a negative control. Distinct nuclear immunostaining for CREB and p-CREB was quantified by an experienced lung cancer pathologist (I.I.W.) under a light microscope. The observer quantified immunohistochemical expression in a blindly fashion regarding the clinical features of the cases (38, 39). In each specimen, up to 1,000 tumor and epithelial cells were examined using a ×20 magnification objective. The intensity of CREB and p-CREB immunostaining was graded on a scale of 0 to 3, with 0 indicating no staining, 1 indicating weak staining, 2 indicating moderate staining, and 3 indicating strong staining. The extent of positive immunoreactivity for CREB and p-CREB (0–1; 1 = 100%) was calculated as the percentage of cells that had nuclear staining for CREB or p-CREB. Scoring of immunohistochemical staining in each specimen was determined as the product of positive immunostaining intensity (0–3) and positive immunoreactivity extent (0–1).

Statistical analysis. A mixed-effect general linear model was used to assess the differences in CREB and p-CREB expression in the normal lung and NSCLC tissue specimens. The Kruskal-Wallis test and Wilcoxon rank sum test were used to assess the relationships between CREB and p-CREB expression in the TMA specimens and patients' demographic and clinicopathologic characteristics. Cox proportional hazards regression models were used to assess the effect of the CREB and p-CREB immunostaining scores on overall survival duration (time from surgery to death of any cause). Survival curves were determined using the Kaplan-Meier product limit estimates, and differences in probability of survival between groups were assessed statistically using the log-rank test. The need for transformation of predictive variables in the Cox proportional hazards regression models was assessed using martingale residual plots. Predictive variables with P values of <0.10 for the univariate Cox proportional hazards regression model were included in a multivariate model. In this multivariate model, backward elimination with a P value cutoff of 0.05 was used; any previously deleted variables were then allowed to reenter the final model if $P < 0.05$.

Results

Higher expression of CREB and p-CREB in NSCLC cell lines than in NHTBE cells. Western blot analysis showed that most of the NSCLC cell lines had higher levels of CREB (9 of 10 cell lines) and p-CREB (7 of 10 cell lines) protein expression than did the NHTBE cells (Fig. 1*A*). qRT-PCR analysis showed a similar pattern in CREB mRNA expression: about 2-fold to 12-fold higher levels in 6 of 10 NSCLC cell lines than in NHTBE cells ($P < 0.05$ in 4 cell lines; $P < 0.01$ in 2 cell lines; Fig. 1*B*); these 6 cell lines also had the highest p-CREB protein expression levels. Unlike CREB protein and mRNA expression, PCR analysis showed that the *CREB* gene copy number was significantly increased in only 2 of the 10 NSCLC cell lines ($P < 0.05$ for both cell lines; Fig. 1*C*). These data clearly show that CREB was overexpressed and highly activated in most of the NSCLC cell lines. Moreover, CREB overexpression in these cell lines mainly resulted from increased *CREB* gene transcription rather than an amplified *CREB* gene copy number.

Higher expression of CREB in frozen NSCLC specimens than in adjacent normal lung tissue specimens. CREB protein was overexpressed (by 50–92%) and more highly activated (by 18–119%) in frozen NSCLC tissue specimens than in adjacent normal tissue specimens according to our Western blot analysis (Fig. 2*A*). Among the six frozen specimens, qRT-PCR analysis showed that CREB mRNA expression was significantly higher in all three squamous cell specimens (by 73–95%) and in two of the three adenocarcinoma specimens (by 65% and 73%) than in the adjacent normal tissue specimens ($P < 0.05$ for all five specimens; Fig. 2*B*). PCR analysis revealed that only one squamous cell carcinoma tissue specimen and one adenocarcinoma tissue specimen had a significantly increased *CREB* gene copy number when compared with normal tissue specimens ($P < 0.05$ for both pairs of specimens; Fig. 2*C*).

The overexpression of CREB and p-CREB in these frozen NSCLC tissue specimens was concordant with our results in the NSCLC cell lines. Taken together, our cell line and frozen tissue specimen results showed that CREB overexpression occurred primarily at the gene transcription level but in some cases could have resulted from an amplified *CREB* gene copy number.

Higher expression of CREB in paraffin-embedded NSCLC specimens than in adjacent normal lung tissue specimens. Immunohistochemical analysis of CREB and p-CREB expression in the 45 whole paraffin-embedded NSCLC specimens and adjacent normal bronchial and bronchiolar epithelial tissue specimens showed stronger nuclear staining for CREB and p-CREB in both adenocarcinoma and squamous cell carcinoma tissue than in normal tissue (Fig. 3*A*). The distributions of the staining scores for the tumor tissue and normal epithelium are shown in Fig. 3*A*. Statistical analysis showed significantly higher immunostaining scores for both CREB (1.37 versus 0.73; $P = 0.013$) and p-CREB (1.96 versus 1.05; $P = 0.0002$) in tumor tissue than in normal tissue.

Association of CREB and p-CREB expression with histologic NSCLC subtype and patients' demographic and clinicopathologic characteristics. We performed immunohistochemical analysis of CREB and p-CREB expression in the 310 NSCLC TMAs to assess potential associations of CREB and p-CREB overexpression with the patients' demographic and clinicopathologic characteristics and histologic subtypes of NSCLC (listed in Supplementary Table S1). Representative images of the staining of CREB and

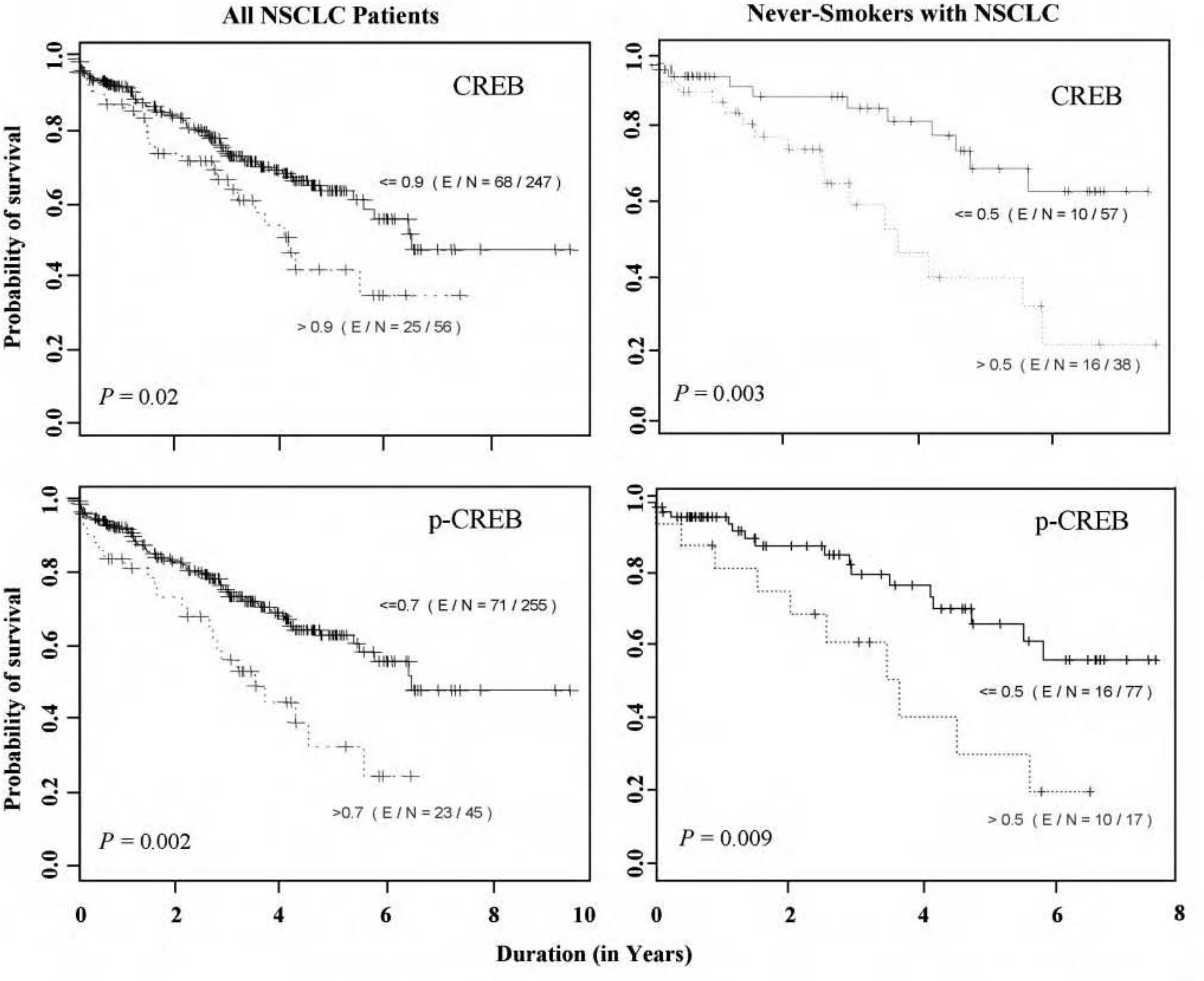


Figure 4. Kaplan-Meier curves of overall survival duration stratified according to CREB (*top*) and p-CREB (*bottom*) expression in the entire cohort (*left*) and in never smokers (*right*). A total of 310 patients had information available for the survival analysis, 94 of whom died. The median overall survival duration was 6.5 y, and the median follow-up duration was 3.2 y. The cutoff points for the CREB and p-CREB immunostaining scores were identified using martingale residual plots with respect to the overall survival duration. The curves are labeled with the corresponding immunostaining scores.

Table 2. Estimation of overall survival durations per smoking status in NSCLC patients from whom tumor TMA specimens were obtained using univariate Cox proportional hazards regression models

Smoking status	Variable estimate	SE	*P*	Hazards ratio
Never				
CREB	0.55	0.24	0.02	1.73
p-CREB	0.99	0.43	0.02	2.69
Former				
CREB	0.24	0.28	0.39	1.28
p-CREB	0.75	0.30	0.01	2.13
Current				
CREB	−0.46	0.56	0.41	0.63
p-CREB	−0.03	0.60	0.96	0.97

p-CREB in the TMA specimens are shown in Fig. 3*B*. According to the Wilcoxon rank sum test, the CREB and p-CREB immunostaining scores were significantly higher in the squamous cell carcinoma specimens than in the adenocarcinoma and BAC specimens (Fig. 3*B*, CREB: 0.61 versus 0.43, $P = 0.002$; p-CREB: 0.32 versus 0.25, $P = 0.008$). The levels of immunohistochemical staining of CREB and p-CREB were lower in the TMA specimens than in the paraffin-embedded whole NSCLC tissue specimens. This phenomenon probably resulted from larger tissue areas and stronger immunostaining in the whole tissue specimens than in the TMA specimens. We detected no significant differences in CREB or p-CREB expression between other demographic and clinicopathologic subpopulations (Table 1).

Effects of CREB and p-CREB overexpression on overall survival duration. The Kaplan-Meier survival curves shown in Fig. 4 show that overexpression of CREB or p-CREB was related to a lower probability of survival. Overexpression of CREB [immunostaining score > 0.9 (E / N = 25/56)] or p-CREB [immunostaining score > 0.7 (E / N = 23/45)] was significantly associated with decreased overall survival duration in patients with NSCLC (P = 0.02 and 0.002, respectively, log-rank test). We also used univariate Cox proportional hazards regression models to determine the effects of covariates on overall survival duration. Factors that significantly affected overall survival were age, histologic subtype, p-CREB expression, pathologic T classification, pathologic N classification, and pathologic tumor-node-metastasis (TNM) stage (Table 1). Multivariate Cox proportional hazards regression model analysis showed that the expression of both CREB and p-CREB was significantly associated with decreased overall survival after accounting for the effects of age and pathologic T and N classification (CREB: $P = 0.02$; p-CREB: $P = 0.01$). One-unit increases in the CREB and p-CREB immunostaining scores increased the risk of death by 51% and 80%, respectively (Table 1).

Effects of CREB and p-CREB overexpression on survival duration according to smoking status. The inverse relationship between the level of CREB and p-CREB expression and the survival duration was significantly dependent on smoking status in the 310 patients from whom the TMA NSCLC specimens were obtained. Overexpression of CREB and p-CREB significantly lowered the probability of survival in never smokers [CREB immunostaining score > 0.5 (E / N = 16/38); p-CREB immunostaining score > 0.5 (E / N = 10/17); Fig. 4, *right*] but not in former or current smokers (data not shown). In addition, univariate Cox proportional hazards regression model analysis showed that both CREB and p-CREB immunostaining scores were significantly inversely correlated with overall survival duration in never smokers, increasing the risk of death by 73% ($P = 0.02$) and 169% ($P = 0.02$), respectively. We also observed this tendency in former smokers, as both CREB and p-CREB immunostaining scores were inversely correlated with overall survival duration, increasing the risk of death by 28% ($P = 0.39$) and 113% ($P = 0.01$), respectively. In contrast, in current smokers, neither CREB nor p-CREB immunostaining scores affected overall survival duration (P = 0.41 and 0.96, respectively; Table 2).

Discussion

Our study of the role of CREB in the development and pathogenesis of NSCLC showed that CREB and p-CREB are overexpressed in patients with NSCLC and that this overexpression is associated with a negative prognosis in never smokers with this disease. CREB expression levels were higher and CREB was more highly activated constitutively in NSCLC cell lines than in NHTBE cells. CREB and p-CREB were also expressed at higher levels in frozen NSCLC tumor specimens than in paired normal tissue specimens. This overexpression of CREB protein resulted primarily from transcriptional overexpression of CREB mRNA rather than from amplification of the *CREB* gene copy number. These results are consistent with a recent report showing that CREB expression is up-regulated at both the protein and mRNA levels in primary acute myeloid leukemia cells compared with that in normal blood cells (9). However, the report indicated that CREB overexpression was associated mainly with an amplified *CREB* gene copy number (in three of four patients with acute myeloid leukemia), whereas we found that CREB overexpression occurred mainly at the gene transcription level (and possibly as a result of an increased *CREB* gene copy number in a few cases). This discrepancy implies mechanistic variation in CREB overexpression in different cancer types. Immunohistochemical analysis of slides containing whole sections of NSCLC and adjacent normal bronchial or bronchiolar epithelial tissue specimens confirmed that expression of both CREB and p-CREB was significantly higher in the tumor specimens than in the normal tissue specimens. In addition, the results of our immunohistochemical study on NSCLC TMAs showed that CREB and p-CREB immunostaining scores were significantly higher in the squamous cell carcinoma specimens than in the adenocarcinoma specimens. Taken together, these results clearly indicate overexpression of CREB in NSCLC.

In addition, the Kaplan-Meier survival curves revealed that overexpression of both CREB and p-CREB was significantly associated with decreased overall survival durations in the 310 patients from whom the TMA NSCLC specimens were obtained. This observed effect of CREB and p-CREB overexpression on survival was mainly contributed by the never-smoker portion of patients. The survival durations in these patients were strongly influenced by the levels of CREB and p-CREB expression, whereas the durations in former and current smokers were less or not affected by them. These data suggest that the expression level of both CREB and p-CREB is a useful biomarker for predicting survival duration, and therefore, CREB could be a therapeutic target for never smokers with NSCLC. Prognosis for and treatment outcome of NSCLC are often in favor of never smokers (40, 41). The reason why overexpression of CREB and p-CREB is a negative prognostic factor in never smokers but not in former or current smokers is unclear. One possible explanation is that survival of cancer cells depends substantially on CREB activity in never smokers, less so in former smokers, and not at all in current smokers. Alternatively, former and current smokers may have other confounding factors that predominate over CREB activity in affecting overall survival.

Cancers in ever smokers may use multiple (proto)oncogenic pathways for growth and survival. Although tobacco smoking is the leading cause of lung cancer, several studies have shown that the biology of lung cancer differs between never smokers and ever smokers [see Sun and colleagues (46) for a comprehensive review; refs. 41–45]. For examples, some well-characterized mutations of EGFR are more frequently detected in a subset of lung cancer patients who are never smokers, whereas mutations in KRas are more prevalent in smokers and such mutation statuses are mutually exclusive in each given set of lung cancers (42, 46). CREB is activated by two downstream pathways of EGFR signaling, namely, Ras-Raf-ERK-RSK (26, 47) and PI3K-Akt (48) pathways, but it is not clear whether these pathways differentially regulate CREB expression and activity. Further studies are required to determine whether CREB plays a differential role in the pathogenesis of lung cancers between smokers and never smokers.

Of note is that CREB may also play a significant role in the progression of lung cancer at its early stage. Studies showed that the transcriptional activity of CREB mediated tobacco smoke–stimulated overexpression of amphiregulin (49), which is associated with poor prognosis in patients with NSCLC, indicating that CREB may play an important role in the early stage of lung carcinogenesis in smokers (50). Recent studies examining the role of CREB in lung cancer have shown an elevated expression of p-CREB in lung tumors generated in insulin-like growth factor II–overexpressing transgenic mice and that CREB played an important role in survival of lung cancer cell lines (51, 52). Further studies are warranted to molecularly characterize the differential role of CREB in smokers and never smokers with NSCLC tumors.

CREB overexpression may lead to up-regulation of genes and activation of signaling pathways that support lung tumor growth and survival. To test this hypothesis, we have gathered data indicating that inactivation of CREB or reduction of CREB expression (via the expression of a dominant repressor of CREB or small interfering RNA against CREB) reduces the expression of antiapoptotic genes and consequently inhibits the growth and survival of NSCLC cells (53). The present study showed that increased expression of CREB and p-CREB correlates with decreased overall survival durations in lung cancer patients, suggesting that sustained overexpression of CREB in malignant cells supports the growth and survival of tumor cells.

In summary, the present study provided important initial insights into the role of CREB in the development and pathogenesis of NSCLC. First, CREB was overexpressed and highly active constitutively in NSCLC cell lines and banked NSCLC specimens. Second, CREB overexpression seemed to be more attributable to increased CREB mRNA transcription than *CREB* gene amplification, suggesting that CREB overexpression is a correctable therapeutic target for patients with NSCLC. Third, overexpression of both CREB and p-CREB was independently correlated with significantly decreased overall survival durations in never smokers but not in current or former smokers with NSCLC. However, further extensive studies are required to determine whether CREB plays a differential role in the development of lung cancer in never and ever smokers. To the best of our knowledge, this study is the first to provide evidence that CREB overexpression and p-CREB overexpression are negative prognostic factors in never smokers with NSCLC. Therefore, targeting CREB, such as with the use of CREB inhibitors, may be a preventive strategy for individuals at high risk for NSCLC and/or a targeted therapeutic strategy for this disease, especially among never smokers.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 9/14/2007; revised 3/19/2008; accepted 4/28/2008.

Grant support: National Heart, Lung, and Blood Institute grant R01-HL-077556 (J.S. Koo), Department of Defense VITAL grant W81XW-04-1-0142 (J.S. Koo), Head and Neck Specialized Program of Research Excellence grant P50 CA097007, Developmental Research Program Grant (J.S. Koo), and National Cancer Institute Cancer Center Support grant CA-16672 (The University of Texas M. D. Anderson Cancer Center).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Drs. Seung-Wook Kim and Jeong-Soo Hong for their technical assistance and Dr. Wen-Cheng Chung and Don Norwood for critical editing of the manuscript.

References

1. Kamath AV, Chhajed PN. Role of bronchoscopy in early diagnosis of lung cancer. Indian J Chest Dis Allied Sci 2008;58:71–96.

2. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2006;56:106–30.

3. Broker LE, Giaccone G. The role of new agents in the treatment of non-small cell lung cancer. Eur J Cancer 2002;38:2347–61.

4. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.

5. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.

6. Sequist LV, Bell DW, Lynch TJ, Haber DA. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. J Clin Oncol 2007;25:587–95.

7. Ionescu AM, Schwarz EM, Vinson C, et al. PTHrP modulates chondrocyte differentiation through AP-1 and CREB signaling. J Biol Chem 2001;276:11639–47.

8. Riccio A, Ahn S, Davenport CM, Blendy JA, Ginty DD. Mediation by a CREB family transcription factor of NGF-dependent survival of sympathetic neurons. Science 1999;286:2358–61.

9. Conkright MD, Montminy M. CREB: the unindicted cancer co-conspirator. Trends Cell Biol 2005;15:457–9.

10. Shankar DB, Sakamoto KM. The role of cyclic-AMP binding protein (CREB) in leukemia cell proliferation and acute leukemias. Leuk Lymphoma 2004;45:265–70.

11. Rudolph D, Tafuri A, Gass P, Hammerling GJ, Arnold B, Schutz G. Impaired fetal T cell development and perinatal lethality in mice lacking the cAMP response element binding protein. Proc Natl Acad Sci U S A 1998;95:4481–6.

12. Long F, Schipani E, Asahara H, Kronenberg H, Montminy M. The CREB family of activators is required for endochondral bone development. Development 2001;128:541–50.

13. Mayr B, Montminy M. Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol 2001;2:599–609.

14. Aggarwal S, Kim SW, Cheon K, Tabassam FH, Yoon JH, Koo JS. Nonclassical action of retinoic acid on the activation of the cAMP response element-binding protein in normal human bronchial epithelial cells. Mol Biol Cell 2006;17:566–75.

15. Song KS, Lee WJ, Chung KC, et al. Interleukin-1β and tumor necrosis factor-α induce MUC5AC overexpression through a mechanism involving ERK/p38 mitogen-activated protein kinases-MSK1-CREB activation in human airway epithelial cells. J Biol Chem 2003;278:23243–50.

16. Cho KN, Choi JY, Kim CH, et al. Prostaglandin E2 induces MUC8 gene expression via a mechanism involving ERK MAPK/RSK1/cAMP response element binding protein activation in human airway epithelial cells. J Biol Chem 2005;280:6676–81.

17. Shaywitz AJ, Greenberg ME. CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. Annu Rev Biochem 1999;68:821–61.

18. Herzig S, Long F, Jhala US, et al. CREB regulates hepatic gluconeogenesis through the coactivator PGC-1. Nature 2001;413:179–83.

19. Koo SH, Flechner L, Qi L, et al. The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. Nature 2005;437:1109–11.

20. Desdouets C, Matesic G, Molina CA, et al. Cell cycle regulation of cyclin A gene expression by the cyclic AMP-responsive transcription factors CREB and CREM. Mol Cell Biol 1995;15:3301–9.

21. Xiang H, Wang J, Boxer LM. The role of the cAMP-response element in the Bcl-2 promoter in the regulation of endogenous Bcl-2 expression and apoptosis in murine B cells. Mol Cell Biol 2006;26:8599–606.

22. McCauslin CS, Heath V, Colangelo AM, et al. CAAT/enhancer-binding protein δ and cAMP-response element-binding protein mediate inducible expression of the nerve growth factor gene in the central nervous system. J Biol Chem 2006;281:17681–8.

23. Oike Y, Takakura N, Hata A, et al. Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. Blood 1999;93:2771–9.

24. Stylianou E, Saklatvala J. Interleukin-1. Int J Biochem Cell Biol 1998;30:1075–9.

25. Nyormoi O, Bar-Eli M. Transcriptional regulation of metastasis-related genes in human melanoma. Clin Exp Metastasis 2003;20:251–63.

26. Kim SW, Hong JS, Ryu SH, Chung WC, Yoon JH, Koo JS. Regulation of mucin gene expression by CREB via a nonclassical retinoic acid signaling pathway. Mol Cell Biol 2007;27:6933–47.

27. Jean D, Harbison M, McConkey DJ, Ronai Z, Bar-Eli M. CREB and its associated proteins act as survival factors for human melanoma cells. J Biol Chem 1998;273:24884–90.

28. Sofi M, Young MJ, Papamakarios T, Simpson ER, Clyne CD. Role of CRE-binding protein (CREB) in aromatase expression in breast adipose. Breast Cancer Res Treat 2003;79:399–407.

29. Crans-Vargas HN, Landaw EM, Bhatia S, Sandusky G, Moore TB, Sakamoto KM. Expression of cyclic adenosine monophosphate response-element binding protein in acute leukemia. Blood 2002;99:2617–9.

30. Shankar DB, Cheng JC, Sakamoto KM. Role of cyclic AMP response element binding protein in human leukemias. Cancer 2005;104:1819–24.

31. Shankar DB, Cheng JC, Kinjo K, et al. The role of CREB as a proto-oncogene in hematopoiesis and in acute myeloid leukemia. Cancer Cell 2005;7:351–62.

32. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.

33. Beasley MB, Brambilla E, Travis WD. The 2004 World Health Organization classification of lung tumors. Semin Roentgenol 2005;40:90–7.

34. Koo JS, Jetten AM, Belloni P, Yoon JH, Kim YD, Nettesheim P. Role of retinoid receptors in the regulation of mucin gene expression by retinoic acid in human tracheobronchial epithelial cells. Biochem J 1999;338:351–7.

35. Koo JS, Yoon JH, Gray T, Norford D, Jetten AM, Nettesheim P. Restoration of the mucous phenotype by retinoic acid in retinoid-deficient human bronchial cell cultures: changes in mucin gene expression. Am J Respir Cell Mol Biol 1999;20:43–52.

36. Kolodziejski PJ, Musial A, Koo JS, Eissa NT. Ubiquitination of inducible nitric oxide synthase is required for its degradation. Proc Natl Acad Sci U S A 2002;99:12315–20.

37. Kim SW, Cheon K, Kim CH, et al. Proteomics-based identification of proteins secreted in apical surface fluid of squamous metaplastic human tracheobronchial epithelial cells cultured by three-dimensional organotypic air-liquid interface method. Cancer Res 2007;67:6565–73.

38. Shen J, Behrens C, Wistuba II, et al. Identification and validation of differences in protein levels in normal, premalignant, and malignant lung cells and tissues using high-throughput Western Array and immunohistochemistry. Cancer Res 2006;66:11194–206.

39. Prudkin L, Behrens C, Liu DD, Zhou X, Ozburn N, Bekele BN. Loss and reduction of Fus1 protein expression is a frequent phenomenon in the pathogenesis of lung cancer. Clin Cancer Res 2008;14:41–7.

40. Nordquist LT, Simon GR, Cantor A, Alberts WM, Bepler G. Improved survival in never-smokers vs current smokers with primary adenocarcinoma of the lung. Chest 2004;126:347–51.

41. Tsao AS, Liu D, Lee JJ, Spitz M, Hong WK. Smoking affects treatment outcome in patients with advanced nonsmall cell lung cancer. Cancer 2006;106:2428–36.

42. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97:339–46.

43. Lam DC, Girard L, Ramirez R, et al. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. Cancer Res 2007;67:4638–47.

44. Toh CK, Gao F, Lim WT, et al. Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. J Clin Oncol 2006;24:2245–51.

45. Gazdar AF, Thun MJ. Lung cancer, smoke exposure, and sex. J Clin Oncol 2007;25:469–71.

46. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. Nat Rev Cancer 2007;7:778–90.

47. Xing J, Ginty DD, Greenberg ME. Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. Science 1996;273:959–63.

48. Du K, Montminy M. CREB is a regulatory target for the protein kinase Akt/PKB. J Biol Chem 1998;273:32377–9.

49. Du B, Altorki NK, Kopelovich L, Subbaramaiah K, Dannenberg AJ. Tobacco smoke stimulates the transcription of amphiregulin in human oral epithelial cells: evidence of a cyclic AMP-responsive element binding protein-dependent mechanism. Cancer Res 2005;65:5982–8.

50. Fontanini G, De Laurentiis M, Vignati S, et al. Evaluation of epidermal growth factor-related growth factors and receptors and of neoangiogenesis in completely resected stage I-IIIA non-small-cell lung cancer: amphiregulin and microvessel count are independent prognostic indicators of survival. Clin Cancer Res 1998;4:241–9.

51. Moorehead RA, Sanchez OH, Baldwin RM, Khokha R. Transgenic overexpression of IGF-II induces spontaneous lung tumors: a model for human lung adenocarcinoma. Oncogene 2003;22:853–7.

52. Linnerth NM, Baldwin M, Campbell C, Brown M, McGowan H, Moorehead RA. IGF-II induces CREB phosphorylation and cell survival in human lung cancer cells. Oncogene 2005;24:7310–9.

53. Aggarwal S, Kim S-W, Ryu S-H, Chung W-C, Koo JS. Growth suppression of lung cancer cells by targeting cyclic AMP response element-binding protein. Cancer Res 2008;68:981–8.

Short Communication

Differential Methylation of a Short CpG-Rich Sequence within Exon 1 of *TCF21* Gene: A Promising Cancer Biomarker Assay

Narayan Shivapurkar,[1,2] Victor Stastny,[1] Yang Xie,[3] Clemens Prinsen,[6] Eugene Frenkel,[4] Bogdan Czerniak,[7] Frederik B. Thunnissen,[6] John D. Minna,[1,4,5] and Adi F. Gazdar[1,2]

[1]Hamon Center for Therapeutic Oncology Research and Departments of [2]Pathology, [3]Clinical Sciences, [4]Internal Medicine, and [5]Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas; [6]Pathology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands; and [7]University of Texas M. D. Anderson Cancer Center, Houston, Texas

Abstract

Detection of cancer cells at early stages could potentially increase survival rates in cancer patients. Aberrant promoter hypermethylation is a major mechanism for silencing tumor suppressor genes in many kinds of human cancers. A recent report from our laboratory described the use of quantitative methylation-specific PCR assays for discriminating patients with lung cancer from those without lung cancer using lung biopsies as well as sputum samples. *TCF21* is known to be essential for differentiation of epithelial cells adjacent to mesenchyme. Using restriction landmark genomic scanning, a recent study identified *TCF21* as candidate tumor suppressor at 6q23-q24 that is epigenetically inactivated in lung and head and neck cancers. Using DNA sequencing technique, we narrowed down a short CpG-rich segment (eight specific CpG sites in the CpG island within exon 1) of the *TCF21* gene, which was unmethylated in normal lung epithelial cells but predominantly methylated in lung cancer cell lines. We specifically targeted this short CpG-rich sequence and developed a quantitative methylation-specific PCR assay suitable for high-throughput analysis. We showed the usefulness of this assay in discriminating patients with lung cancer from those without lung cancer using biopsies and sputum samples. We further showed similar applications with multiple other malignancies. Our assay might have important implications in early detection and surveillance of multiple malignancies. (Cancer Epidemiol Biomarkers Prev 2008;17(4):995–1000)

Introduction

Transcriptional inactivation of CpG island-containing promoters of tumor suppressor genes by DNA hypermethylation has been well documented in many human cancers (1). Methylation of specific CpG residues within a CpG island of a tumor suppressor gene may reflect gene silencing and indicate, at least in part, the expression status of the gene. Gene promoter hypermethylation potentially provides a noninvasive screen for early cancer detection (2).

Methylation analyses have been conducted using conventional methodologies such as COBRA, direct sequencing, or methylation-specific PCR of the bisulfite-treated DNA. However, these methods are labor intensive, amenable to false-positive results, and not suitable for high-throughput analysis. Methyl Light assays are not only highly specific, sensitive, and reproducible but also are nonsubjective and allow for rapid analysis of many samples at multiple gene loci (3, 4). Recent publications have shown the presence of promoter hypermethylation of various genes in clinical specimens containing exfoliated tumor cells (such as malignant effusions, sputum, serum, etc.; refs. 5-8). Recently, we reported quantitative methylation-specific PCR analysis of sputum DNA based on a panel of methylated genes (9). The panel separated patients with lung cancer from those without lung cancer, showing the potential of the quantitative methylation-specific PCR analysis of sputum as an effective biomarker assay. Subsequently, we decided to explore additional novel markers that might further improve this assay.

TCF21 is known to be essential for differentiation of epithelial cells adjacent to mesenchyme (10). Using restriction landmark genomic scanning, Smith et al. (11) identified *TCF21* as a candidate tumor suppressor at 6q23-q24 that is epigenetically inactivated in lung and head and neck cancers. In our article, using DNA sequencing, we examined lung cancer and bronchial epithelial cell lines for methylation of the CpG island

---

Received 11/28/07; revised 1/31/08; accepted 2/4/08.

Grant support: Early Detection Research Network grant U01CA084971, National Cancer Institute University of Texas SPORE in Lung Cancer grant P50CA70907, and Department of Defense grant W81XWH-04-1-0142.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Current address for F.B. Thunnissen: VU Medical Center, Amsterdam, The Netherlands.

Requests for reprints: Narayan Shivapurkar, Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390. Phone: 214-648-1456; Fax: 214-648-4940. E-mail: Narayan.Shivapurkar@UTSouthwestern.edu

Copyright © 2008 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-07-2808

within exon 1 of the *TCF21* gene (GenBank accession no. AF047419). We identified a short CpG-rich segment that was unmethylated in normal bronchial cells but predominantly methylated in lung cancer cell lines. We decided to develop a high-throughput quantitative bioassay to determine prevalence of methylation in clinical samples and evaluate its potential as cancer biomarker assay applicable for multiple cancer types.

Materials and Methods

Surgically resected non-small lung and bladder cancers and their adjacent nonmalignant tissues were obtained from the University of Texas M. D. Anderson Cancer Center. Surgically resected breast cancers and their adjacent nonmalignant breast tissues were obtained from Parkland Memorial Hospital. Leukemia cases, all acute myeloid leukemia, were obtained from Parkland Memorial Hospital. Peripheral blood mononuclear cells were obtained from healthy individuals with a family history of cancer. We established all cell lines used in this study. Sputum samples were obtained from 13 patients with non–small cell lung cancer (NSCLC) and 25 individuals with chronic obstructive pulmonary disease; all were heavy smokers without lung cancer in the Canisius Wilhelmina Hospital. Three-day pooled early morning sputum samples were collected in Saccomanno's fixative (2% polyethylene glycol in 50% ethanol). Informed consent and institutional review board permission were obtained at each site.

Gene Expression in Cell Lines. Gene expression studies were conducted as described previously (11) with some modifications. Semiquantitative real-time PCR was carried out by using QuantiTect SYBR Green PCR kit. The expression levels were quantitated using comparative Ct method. In case of both *TCF21* and *WNT4*, means of expression values for the two human bronchial epithelial cells (HBEC) were considered to have a value of 1. 5-Aza-2-deoxycytidine treatment of lung cell lines (H1299, H2887, HCC95, and H661) was done using protocol as described previously (12).

DNA Extraction and Bisulfite Modification and DNA Sequencing. Genomic DNA was extracted from cell lines, primary tumors and nonmalignant cells, sputum samples as described previously (9). Sodium bisulfite treatment was done as described previously (13). The modified DNA was used as a template for quantitative PCR (qPCR) analysis. The DNA sequencing was carried out as described previously using Applied Biosystems prism dye terminator cycle sequencing method (Perkin-Elmer; ref. 14). Primers were designed to exclude CG sites, rendering DNA amplification independent of the methylation status: forward 5′-ATGTGGAGGATTTTAAGAGGT-3′ and reverse 5′-CTAAAAAAAACCTTACTCAACACTC-3′. The sequences were confirmed by sequencing in both directions.

Quantitative Real-time PCR Analysis. qPCR analysis was done using the Chromo4 MJ Research Real-time PCR System. Sodium bisulfite–treated genomic DNA was amplified by fluorescence-based real-time methylation-specific PCR using TaqMan technology as described previously (9, 15, 16). In brief, primers (forward 5′-CGAGGAGAGTTTTAATTGCGAGA-3′ and reverse 5′-CCTAACTAACCCCGCTCAAAAAA-3′) and probe (5′-FAM-TAGAAGGGTCGCGGCGGTTTGGBHQ-1-3′) were designed to specifically amplify bisulfite-converted DNA within the region of the test genes that was differentially methylated between expression-positive and expression-negative cell lines (Fig. 1). The nonmethylated form of *MYOD1* was used as an internal reference standard (9). The sputum DNA samples were coded and shipped from The Netherlands and analyzed in Dallas in a blinded fashion. Some of the lung cancer DNA was also independently analyzed by three investigators (N.S., V.S., and C.P.) in a blinded fashion. The data from all the three investigators were in close agreement (data not shown).

Nested qPCR protocol was accomplished by doing two rounds of PCR as reported previously (17). In the first round, external primers used for DNA sequencing were used. The successful amplification in the first round was confirmed through agarose gel electrophoresis. The second round was carried on the 500× dilution of the first round reaction using the probe and primer sets as in quantitative methylation-specific PCR protocol described above.

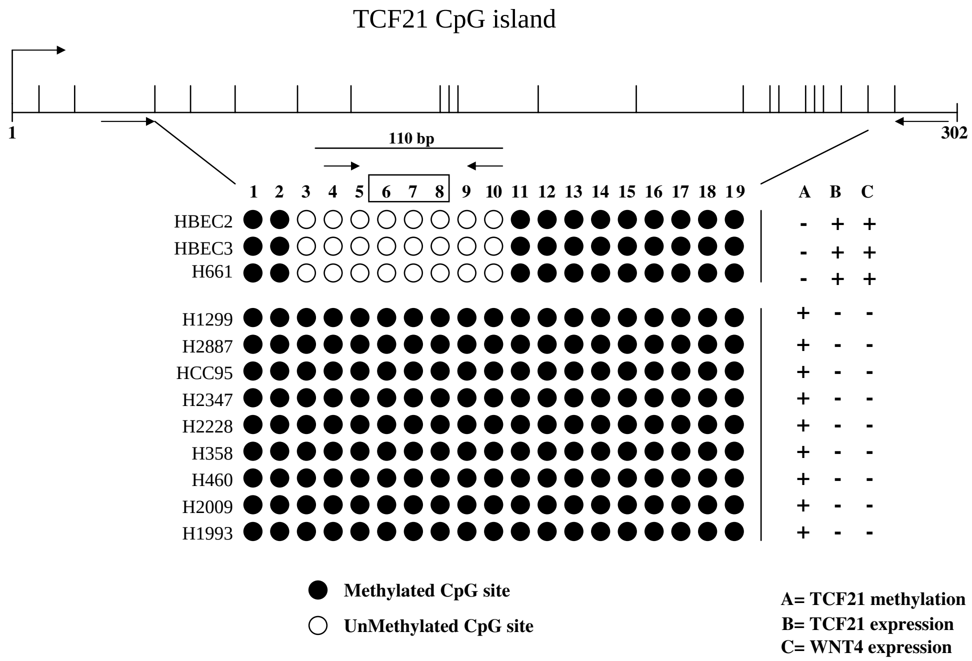
Statistical Analysis. The receiver operating characteristic curves, a plot of the sensitivity versus specificity across all possible cutoff values, were used to identify the accuracy of a marker in discriminating cancer from nonmalignant tissue. The quantitative methylation data for the gene were correlated with tumor stage using the Mann-Whitney *U* test, which does not require parametric assumption on the distribution of quantitative methylation. Statistical differences between groups were examined using Fisher's exact test. P values <0.05 were considered significant.

Results

We sequenced the CpG island within exon 1 of the *TCF21* gene from 10 lung cancer cell lines and 2 HBECs (18) and also analyzed them for gene expression. The results showed that all the 19 CpG sites were methylated in 9 of 10 lung cancer cell lines, whereas 8 specific CpG sites of the 19 were unmethylated only in H661 (a cancer cell line) and in the HBECs (Fig. 1). All methylation-positive cancer cell lines were negative for *TCF21* expression (relative expression <0.50). H661 and HBECs were negative for methylation and positive for *TCF21* expression (relative expression ≥1; Fig. 1). 5-Aza-2′-deoxycytidine treatment to four lung cell lines (H1299, H2887, HCC95, and H661) resulted into reactivation of *TCF21* expression in all the cell lines, except H661, which was positive for *TCF21* expression in untreated population (data not shown). The results suggest that methylation of the specific sites is related to gene silencing. However, it is possible that other factors in addition to methylation (19) might also contribute to gene silencing in these lung cancer cell lines. We examined *WNT4* expression for two reasons. Our preliminary studies showed consistent loss of its expression in lung cancer cell lines compared with HBECs.[8] Secondly, the gene was

---

[8] Unpublished data.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	A	B	C
HBEC2	●	●	○	○	○	○	○	○	○	○	●	●	●	●	●	●	●	●	●	-	+	+
HBEC3	●	●	○	○	○	○	○	○	○	○	●	●	●	●	●	●	●	●	●	-	+	+
H661	●	●	○	○	○	○	○	○	○	○	●	●	●	●	●	●	●	●	●	-	+	+
H1299	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H2887	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
HCC95	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H2347	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H2228	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H358	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H460	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H2009	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-
H1993	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	-	-

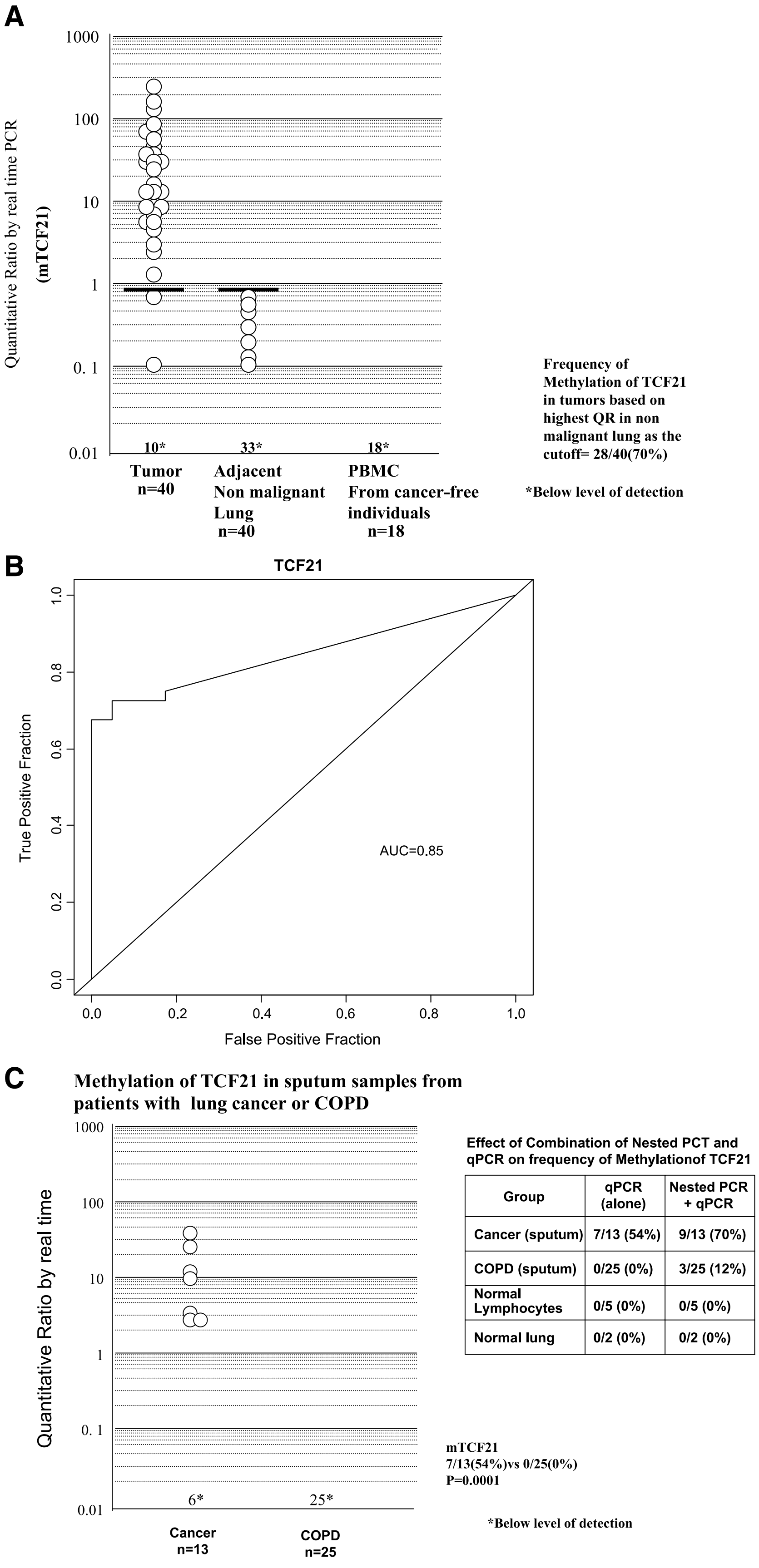
Figure 1. Direct sequencing of the bisulfite PCR product from the CpG island within exon 1 of the *TCF21* gene for 10 lung cancer cell lines and 2 normal lung epithelial cell lines (HBEC2 and HBEC3). *Closed circles,* methylated CpG sites; *open circles,* unmethylated CpG sites. The expression levels were quantitated using comparative Ct method. In both *TCF21* and *WNT4* expression analysis, means of expression values for the two HBECs were considered to have a value of 1. (expression positive, relative expression >1; expression negative, relative expression <0.50). Sequencing and expression analyses were conducted as described in Materials and Methods. *Arrows,* locations of primers; *open box,* location of the probe.

also shown to be induced significantly when the lung cancer cells were transfected with *TCF21* (11). Interestingly, all the lung cancer cell lines (except H661) were negative for *WNT4* expression (relative expression <0.50; Fig. 1). Based on all above observations, the specific CpG region was considered a potential target to develop an assay to discriminate between cancer cells and normal cells.

We tested the qPCR assay for methylation of *TCF21* by analysis of 40 NSCLC tumors (22 adenocarcinoma and 18 squamous carcinomas) and their adjacent nonmalignant lung tissues. Additionally, peripheral blood mononuclear cells from 12 healthy subjects recruited for genetic epidemiology studies were also quantitatively analyzed. Figure 2A shows the quantitative methylation data for NSCLC and their adjacent nonmalignant lung tissue. We found aberrant methylation of *TCF21* in 30 of 40 (75%) of primary lung tumors (QR, 0-129.67) and in 7 of 40 (18%) adjacent lung tissue (QR, 0-0.8). Based on the highest QR in nonmalignant tissue as the cutoff, 28 of 40 (70%) of primary NSCLC were found to have *TCF21* methylated. Additionally, peripheral blood mononuclear cells from 18 healthy subjects recruited for genetic epidemiologic studies were analyzed for *TCF21* methylation and found to be below the levels of detection. Figure 2B shows the receiver operating characteristic curve, which provides evidence for the excellent discriminatory capacity of *TCF21* methylation in separating cancer from adjacent normal lung tissue. We also correlated quantitative methylation data for *TCF21* with tumor stage using the Mann-Whitney *P* test. Increased methylation levels (QR) were observed with increase in stage (stage I-II, $P = 0.008$; stage I combined with stages II and III, $P = 0.002$).

We further tested the qPCR assay for methylation of *TCF21* in sputum DNA from patients with lung cancer (Fig. 2C). We analyzed 38 sputum DNA samples (13 from cancer and 25 noncancer patients) for methylation of *TCF21*. We found 7 of 13 (54%) cancer sputa (QR, 0-34.63) and 0 of 25 (0%) noncancer sputa showed methylation of *TCF21* ($P < 0.0001$; Fig. 2C). Combination of nested PCR and qPCR led to increase in methylation frequencies from 7 of 13 (54%) to 9 of 13 (70%) in cancer cases and from 0 of 25 (0%) to 3 of 25 (12%) in chronic obstructive pulmonary diseases (table in Fig. 2C). Whereas no methylation was detected (Ct values of 50) in normal lung and normal lymphocytes, the methylated status of the newly observed five positive samples (all with Ct values <35) following two-round PCR protocol was confirmed through DNA sequencing (data not shown). We speculate that these five cases might be at higher risk to develop lung cancer than other cases, all of which (like normal lung and normal lymphocytes) showed Ct values of 50 even after two rounds of PCR. We have proven the feasibility of the nested qPCR combination approach for increasing the sensitivity without significantly compromising the specificity. Further development of the nested qPCR assay is in progress to assess real effect of this approach.

We further tested the qPCR assay for methylation of *TCF21* by analysis of 30 breast tumors and their adjacent nonmalignant tissues. We found aberrant methylation of



Effect of Combination of Nested PCT and qPCR on frequency of Methylationof TCF21

Group	qPCR (alone)	Nested PCR + qPCR
Cancer (sputum)	7/13 (54%)	9/13 (70%)
COPD (sputum)	0/25 (0%)	3/25 (12%)
Normal Lymphocytes	0/5 (0%)	0/5 (0%)
Normal lung	0/2 (0%)	0/2 (0%)

Figure 2. **A.** Methylation levels of m*TCF21* in NSCLC (*Tumor*), adjacent nonmalignant lung, and peripheral blood mononuclear cells from cancer-free individuals. Methylation levels were quantitated by semiquantitative real-time PCR. Real-time analysis was done as described in Materials and Methods. Quantitative ratio is defined as the ratio of the fluorescence emission intensity values for the PCR products of the biomarker gene to those of PCR products of *MYOD1* multiplied by 100. The ratio is a measure for the relative level of methylation in an individual sample. Because values are expressed on a log scale, completely negative values are expressed as values of 0.01. *Solid horizontal bar,* threshold above which the samples are considered positive for methylation. **B.** Receiver operating characteristic curves for m*TCF21* at separating cancer from the adjacent nonmalignant lung. Receiver operating characteristic curves are plots of the true-positive rate (*Y axis*) against the false-positive rate (*X axis*) for the different possible cutoff points of a diagnostic test. The closer the curve follows the left-hand border and then the top border of the receiver operating characteristic space, the more accurate the test [area under the curve (AUC) = 0.85]. **C.** Methylation levels of m*TCF21* in sputum from NSCLC patients and in sputa from patients without malignancy (chronic obstructive pulmonary diseases). Methylation levels were quantitated as described in Materials and Methods. The table shows the frequency of methylation in the sputum samples following the combination of nested PCR and qPCR. The two-round PCR protocol was carried out as described in Materials and Methods.

Table 1. Quantitative data for methylation of *TCF21* in multiple malignancies

Tissue	Quantitative ratio (range)		% Samples positive for methylation in tumors with QR above the highest in normal group
	Cancer	Normal	
Breast	0-850.68 (*n* = 30)	0-2.0 (*n* = 30)	20/30 (67)
Bladder	0-52.60 (*n* = 41)	0-0 (*n* = 6)	18/41 (44)
Colon	0-256.83 (*n* = 20)	0-8.12 (*n* = 20)	15/20 (75)
Leukemia	0-74.80 (*n* = 27)	0-0 (*n* = 12)	11/27 (41)

TCF21 in 26 of 30 (87%) of breast tumors (QR, 0-850.68) and in 13 of 30 (40%) adjacent breast tissue (QR, 0-2; Table 1). Based on the highest QR in nonmalignant tissue as the cutoff, 20 of 30 (67%) of primary breast tumors were found to have *TCF21* methylated. The data show that *TCF21* methylation has excellent discriminatory capacity at separating cancer from adjacent nonmalignant breast tissue (Table 1).

Additionally, we tested the qPCR assay for methylation of *TCF21* in bladder cancers, colon cancers, and leukemias with corresponding nonmalignant tissues (Table 1). Also, in the case of bladder cancer cases, a strong trend was observed between methylation and invasiveness (52% in invasive and 27% in noninvasive cases). In all the malignancies, *TCF21* methylation appears to have excellent discriminatory capacity at separating cancer from nonmalignant tissue.

Discussion

In this article, we identified a short CpG-rich sequence within exon 1 of the *TCF21* gene that was differentially methylated in lung cancer cells compared with normal lung cells. We targeted this region to design a methylation-specific qPCR assay suitable for high-throughput analysis. We successfully showed the application of this assay, in discriminating patients with lung cancer from those without lung cancer, using lung cancer biopsies and sputa from cancer patients. We further showed similar application of this assay with multiple other malignancies, covering different organ systems.

Transition between epithelial cell to mesenchymal cell is known to occur during tumorigenesis (20, 21). Epithelial-mesenchymal transition has been described in many cancers and correlates with clinical outcome (20). Malignant lesions are often defined by their differentiation status, where benign tumors and low-grade cancers typically retain their epithelial phenotype and malignant cells acquire more fibroblastic mesenchymal phenotype (20). As mentioned in the results, Smith et al. (11) previously reported highly significant induction of *WNT4* expression in lung cancer cells (*TCF21* expression negative) when transfected with *TCF21*. *WNT4* has been associated previously with epithelial-mesenchymal transition and epithelial phenotype (22). We observed significant expression of *WNT4* in HBECs with down-regulation in all the lung cancer cells analyzed in our study, except H661 (*TCF21* expression positive). Our results suggest the interesting possibility that *WNT4* might be one of the proximate downstream targets of *TCF21*. This possibility should be further explored. Thus, our bioassay to quantitate the prevalence of methylation of *TCF21* in clinical samples may help determine the risk of cancer progression in patients and have considerable clinical application.

It appears from our data that methylation of the short CpG-rich sequence in the *TCF21* gene may have significant potential as a biomarker assay in lung cancer and multiple other cancers. As evident from sputum analysis, the feasibility of increasing the sensitivity without significantly compromising the specificity (through combination of nested and qPCR) might make the assay more appealing for clinical as well as epidemiologic applications. Studies are in progress to further evaluate the potential of the assay in noninvasive detection of lung and other cancers.

References

1. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999;21:163–7.
2. Sidransky D. Emerging molecular markers of cancer. Nat Rev Cancer 2002;2:210–219.
3. Eads CA, Nickel AE, Laird PW. Complete genetic suppression of polyp formation and reduction of CpG-island hypermethylation in Apc(Min/+) Dnmt1-hypomorphic mice. Cancer Res 2002;62:1296–9.
4. Eads CA, Lord RV, Wickramasinghe K, et al. Epigenetic patterns in the progression of esophageal adenocarcinoma. Cancer Res 2001;61:3410–8.
5. Brabender J, Usadel H, Danenberg KD, et al. Adenomatous polyposis coli gene promoter hypermethylation in non-small cell lung cancer is associated with survival. Oncogene 2001;20:3528–32.
6. Usadel H, Brabender J, Danenberg KD, et al. Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. Cancer Res 2002;62:371–5.
7. Harden SV, Tokumaru Y, Westra WH, et al. Gene promoter hypermethylation in tumors and lymph nodes of stage I lung cancer patients. Clin Cancer Res 2003;9:1370–5.
8. Topaloglu O, Hoque MO, Tokumaru Y, et al. Detection of promoter hypermethylation of multiple genes in the tumor and bronchoalveolar lavage of patients with lung cancer. Clin Cancer Res 2004;10:2284–8.
9. Shivapurkar N, Stastny V, Suzuki M, et al. Application of a methylation gene panel by quantitative PCR for lung cancers. Cancer Lett 2007;247:56–71.
10. Quaggin SE, Schwartz L, Cui S, et al. The basic-helix-loop-helix protein pod1 is critically important for kidney and lung organogenesis. Development 1999;126:5771–83.
11. Smith LT, Lin M, Brena RM, et al. Epigenetic regulation of the tumor suppressor gene TCF21 on 6q23-24 in lung and head and neck cancer. Proc Natl Acad Sci U S A 2006;103:982–7.
12. Shigematsu H, Suzuki M, Takahashi T, et al. Aberrant methylation of HIN-1 (high in normal-1) is a frequent event in many human malignancies. Int J Cancer 2005;113:600–4.
13. Herman JG, Graff JR, Myohanen S, et al. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–6.
14. Shigematsu H, Takahashi T, Nomura M, et al. Somatic mutations of the HER2 kinase domain in lung adenocarcinomas. Cancer Res 2005;65:1642–6.
15. Toyooka KO, Toyooka S, Maitra A, et al. Establishment and validation of real-time polymerase chain reaction method for CDH1 promoter methylation. Am J Pathol 2002;161:629–34.

16. Shivapurkar N, Stastny V, Takahashi T, et al. Novel real-time PCR assay using a universal molecular marker for diagnosis of hematologic cancers. Int J Cancer 2005;116:656 – 60.

17. Fackler MJ, McVeigh M, Mehrotra J, et al. Quantitative multiplex methylation-specific PCR assay for the detection of promoter hypermethylation in multiple genes in breast cancer. Cancer Res 2004;64:4442 – 52.

18. Ramirez RD, Sheridan S, Girard L, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res 2004;64:9027 – 34.

19. Fuks F. DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev 2005;15:490 – 5.

20. Hugo H, Ackland ML, Blick T, et al. Epithelial-mesenchymal and mesenchymal-epithelial transitions in carcinoma progression. J Cell Physiol 2007;213:374 – 83.

21. Guarino M, Rubino B, Ballabio G. The role of epithelial-mesenchymal transition in cancer pathology. Pathology 2007;39:305 – 18.

22. Taki M, Kamata N, Yokoyama K, et al. Down-regulation of Wnt-4 and up-regulation of Wnt-5a expression by epithelial-mesenchymal transition in human squamous carcinoma cells. Cancer Sci 2003;94:593 – 7.

Cancer **Prevention** Research

Cyclic AMP-Responsive Element Binding Protein– and Nuclear Factor-κB–Regulated CXC Chemokine Gene Expression in Lung Carcinogenesis

Hongxia Sun, Wen-Cheng Chung, Seung-Hee Ryu, Zhenlin Ju, Hai T. Tran, Edward Kim,
Jonathan M. Kurie and Ja Seok Koo

Abstract The recognition of the importance of angiogenesis in tumor progression has led to the development of antiangiogenesis as a new strategy for cancer treatment and prevention. By modulating tumor microenvironment and inducing angiogenesis, the proinflammatory cytokine interleukine (IL)-1β has been reported to promote tumor development. However, the factors mediating IL-1β–induced angiogenesis in non–small cell lung cancer (NSCLC) and the regulation of these angiogenic factors by IL-1β are less clear. Here, we report that IL-1β up-regulated an array of proangiogenic CXC chemokine genes in the NSCLC cell line A549 and in normal human tracheobronchial epithelium cells, as determined by microarray analysis. Further analysis revealed that IL-1β induced much higher protein levels of CXC chemokines in NSCLC cells than in normal human tracheobronchial epithelium cells. Conditioned medium from IL-1β–treated A549 cells markedly increased endothelial cell migration, which was suppressed by neutralizing antibodies against CXCL5 and CXCR2. We also found that IL-1β–induced CXC chemokine gene overexpression in NSCLC cells was abrogated with the knockdown of cyclic AMP-responsive element binding protein (CREB) or nuclear factor κB (NF-κB). Moreover, the expression of the CXC chemokine genes as well as CREB and NF-κB activities was greatly increased in the tumorigenic NSCLC cell line compared with normal, premalignant immortalized or nontumorigenic cell lines. A disruptor of the interaction between CREB-binding protein and transcription factors such as CREB and NF-κB, 2-naphthol-AS-E-phosphate (KG-501), inhibited IL-1β–induced CXC chemokine gene expression and angiogenic activity in NSCLC. We propose that targeting CREB or NF-κB using small-molecule inhibitors, such as KG-501, holds promise as a preventive and/or therapeutic approach for NSCLC.

Lung cancer is the leading cause of cancer deaths both in the United States and worldwide (1). The growth and development of lung cancer as well as of other solid tumors is critically dependent on a functional vascular supply. Numerous lines of evidence have shown that angiogenesis, the formation of new blood vessels from the preexisting vasculature, is one of the critical steps in the entire process of cancer development from tumor growth to distant metastasis (2–5). In addition, a study showed that a solid tumor would remain dormant at a volume of only 2 to 3 mm^3 in the absence of neovascularization and would be unable to metastasize (6). For solid tumors to develop and metastasize, they need to secrete a number of proangiogenic factors to induce the formation of new blood vessels to overcome the physical limitations on the diffusion of nutrients and oxygen within the tumor—a process known as angiogenic switch (3). Recognizing the importance of angiogenesis in the growth of tumors has led to the development of antiangiogenesis as a new strategy for cancer therapy and prevention, a novel concept termed "angioprevention" (7–9). Actually, a series of molecules proposed as chemopreventive agents have been shown to have potent antiangiogenic properties when tested in *in vitro* and *in vivo* angiogenesis models (7).

Angiogenesis can be regulated by various growth factors and cytokines, including vascular endothelial growth factor

---

Authors' Affiliation: Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Received 12/08/2007; revised 06/10/2008; accepted 06/24/2008.

Grant support: Department of Defense VITAL grant W81XW-04-1-0142 (J.S. Koo); National Heart, Lung, and Blood Institute Grant R01-HL-077556 (J.S. Koo); M. D. Anderson Cancer Center SPORE in Head and Neck Cancer (P50 CA097007) (Principal Investigator: S.M. Lippman) Developmental Research Program Grant P50 CA097007 (J.S. Koo); and National Cancer Institute Cancer Center Support Grant CA-16672 to The University of Texas M. D. Anderson Cancer Center.

Note: H. Sun and W-C. Chung contributed equally to this work.

Current address for Z. Ju: Department of Bioinformatics and Computational Biology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Requests for reprints: Ja Seok Koo, Department of Thoracic/Head and Neck Medical Oncology, Unit 432, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-6363; Fax: 713-794-5997; E-mail: jskoo@mdanderson.org.

©2008 American Association for Cancer Research.

doi:10.1158/1940-6207.CAPR-07-0002

(VEGF), basic fibroblast growth factor, transforming growth factors α and β, platelet-derived endothelial cell growth factors, chemokines, and interleukine (IL)-1β (10–14). Recent studies have shown the importance of the tumor microenvironment in facilitating angiogenesis and promoting tumor invasion and metastasis (15–19). Once a tumor is vascularized, the tumor-associated antigens can be recognized by the immune system and the tumor is infiltrated by leukocytes. Although leukocyte infiltration in tumors is often considered to be associated with better prognosis and overall survival, studies have also shown that inflammatory cells can promote tumor cell proliferation, angiogenesis, metastasis, and, hence, tumor development (15, 16). Leukocyte infiltration can influence angiogenesis in tumors because some subsets of leukocytes, especially the tumor-associated macrophages, can secrete both angiostatic and angiogenic factors (17, 18). IL-1 is a proinflammatory cytokine produced mainly by monocytes and macrophages. There are two IL-1 agonistic proteins, IL-1α and IL-1β. IL-1α is a precursor or membrane-associated molecule and is primarily a regulator of intracellular events and a mediator of local reactions. On the other hand, IL-1β acts as a systemic, hormone-like mediator and is only active in a secreted mature form. However, once these two proteins bind to their receptors, they have similar biological activities (20). Both IL-1α and IL-1β can promote tumor angiogenesis, but the role of IL-1β is more evident (14). IL-1 has been shown to contribute to the production of proangiogenic factors VEGF, hepatocyte growth factor, tumor necrosis factor, and CXC chemokines (14, 21). Members of a subfamily of CXC chemokines sharing a characteristic glutamate-leucine-arginine (ELR) motif near the NH_2 terminus of the molecule are chemoattractants for neutrophils and are important for wound repair. The ELR-positive chemokines, including CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, and CXCL8, are proangiogenic, whereas members of another subfamily lacking the ELR motif—ELR-negative chemokines, such as CXCL4, CXCL9, CXCL10, and CXCL11—are in general IFN inducible and are potential inhibitors of angiogenesis. Generally, CXCR2 is the receptor for angiogenic CXC chemokine–mediated angiogenesis, and CXCR3 is the receptor for angiostatic IFN-inducible CXC chemokine inhibition of angiogenesis (13). CXC chemokine ligands and receptors have been shown to play important roles in mediating non–small cell lung cancer (NSCLC)–associated angiogenesis and organ-specific metastases (13). Recently, it has been reported that CXCL5 and CXCL8 protein levels were elevated in tumor specimens freshly isolated from patients with NSCLC and that these two ELR-positive CXC chemokines are important mediators of angiogenesis during NSCLC tumorigenesis (22, 23). Compared with CXCL8, CXCL5 was reported to have a higher degree of correlation with NSCLC-derived angiogenesis (23). In a model system of human NSCLC tumorigenesis in severe combined immunodeficiency mice, CXCL5 expression was found to be directly correlated with tumor growth, tumor-derived angiogenesis, and metastatic potential. Depletion of CXCL5 in this model system resulted in attenuation of both tumor growth and spontaneous metastasis due to the inhibition of angiogenesis (23).

Being a product of tumor infiltrated macrophages, IL-1β is known to increase angiogenesis. However, in NSCLC, what angiogenic factors are induced by IL-1β and how they are regulated by IL-1β are still not clear. To elucidate these critical issues, we conducted a microarray analysis to determine the effect of IL-1β on global gene expression in the NSCLC adenocarcinoma cell line A549 and in normal human tracheobronchial epithelium (NHTBE) cells. We found that IL-1β dramatically induced the expression of an array of proangiogenic CXC chemokine genes and significantly augmented the angiogenic activity of NSCLC. In addition, we found that transcription factors cyclic AMP-responsive element binding protein (CREB) and nuclear factor κB (NF-κB) both play critical roles in the regulation of IL-1β–induced CXC chemokine gene expression and angiogenic activity.

Materials and Methods

Cell cultures, chemicals, and conditioned media

NHTBE cells were purchased from Cambrex and cultured in six-well plates as described previously (24–28). Human umbilical vein endothelial cells (HUVEC) were purchased from Cambrex and maintained in EGM complete endothelial growth medium. HUVECs were used at passage 4. The human NSCLC cell lines A549, H1734, H226, and H2170 were obtained from the American Type Culture Collection and grown in RPMI 1640 containing 10% fetal bovine serum, 100 units/mL penicillin, and 100 μg/mL streptomycin. The BEAS-2B, 1799, 1198, and 1170-I human bronchial epithelial cell lines were obtained from Dr. R. Lotan (The University of Texas M.D. Anderson Cancer Center, Houston, TX) and Dr. A. Klein-Szanto (Fox Chase Cancer Center, Philadelphia, PA) and grown in Keratinocyte Serum-Free Medium (Life Technologies, Inc.) containing epidermal growth factor and bovine pituitary extract (29). All of the cells were cultured at 37°C in a humidified atmosphere of 95% air and 5% CO_2. Conditioned media (CM) were generated as follows: NSCLC cells (3×10^5/mL) were cultured in RPMI 1640 containing 0.5% fetal bovine serum with or without IL-1β and/or 2-naphthol-AS-E-phosphate (KG-501) at different concentrations for 24 h. Cell-free supernatants were collected and stored at −70°C until use.

Reagents and antibodies

IL-1β, Quantikine CXCL5 and CXCL8 ELISA kits, and neutralizing antibodies against CXCL5, CXCL8, CXCR2, and VEGF were purchased from R&D Systems. KG-501 was purchased from Sigma-Aldrich and dissolved in DMSO. Antibodies against NF-κB p65, CREB, and phospho-CREB (Ser^{133}) were purchased from Santa Cruz Biotechnology, Inc., and Upstate, respectively. A monoclonal antibody (mAb) against β-actin was from Sigma-Aldrich. Transwell chambers with polyethylene terephthalate membranes containing 8-μm pores were obtained from BD Biosciences.

Microarray analysis

After confluence, NHTBE and A549 cells were treated with control medium or the same medium containing 2.5 ng/mL IL-1β for 8 h before total RNA extraction. Total RNA was isolated with the RNeasy Mini Kit (Qiagen). The integrity of mRNA and the relative rRNA contamination were analyzed with the RNA 6000 Nano LabChip (Agilent Technologies) and the Agilent 2100 bioanalyzer (Agilent Technologies). The RNA from control group was amplified and labeled with cyanine 3, and that from IL-1β–treated cells with cyanine 5. Equal amounts of the differently labeled RNAs were then mixed and hybridized with 44K whole human genome oligonucleotide microarrays (Agilent Technologies). After hybridization, the arrays were scanned and the resulting images were analyzed using the Agilent feature extraction software program (GE2, version 5.91, Agilent Technologies).

Quantitative reverse transcription-PCR

Validation of the differentially expressed genes in NHTBE and NSCLC cells was done using an iCycler real-time PCR detection system (Bio-Rad) with gene-specific primers. Primers for human glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), the reference gene, were 5′-TGCACCACCAACTGCTTAGC (forward) and 5′-GGCATGGACTGTGGTCATGAG (reverse); for *CXCL1*, 5′-AGTGACAAATCCAACTGACC (forward) and 5′-GATGCTCAAACACATTAGGC (reverse); for *CXCL2*, 5′-CCCAAGTTAGTTCAATCCTG (forward) and 5′-TTCCTCAGCCTCTATCACAG (reverse); for *CXCL3*, 5′-CTTGTCTCAACCCCGCATCC (forward) and 5′-TCTGGTAAGGGCAGGGACCA (reverse); for *CXCL5*, 5′-TCCAATCTCCGCTCCTCCAC (forward) and 5′-AGCAGCAGCAGCACCAACAG (reverse); for *CXCL6*, 5′-GTTTGTCTGGACCCGGAAGC (forward) and 5′-TCCGCTGAAGACTGGGCAAT (reverse); for *CXCL8*, 5′-GCATAAAGACATACTCCAAACC (forward) and 5′-ACTTCTCCACAACCCTCTG (reverse); and for *CREB*, 5′-AAGCTGAAAACCAACAAATGACAGTT (forward) and 5′-TGAACTGTCTGCCCATTGG (reverse). Single-stranded cDNAs were synthesized in 50 μL of reverse transcription (RT) mix containing 1 μg of total RNA using the GeneAmp RNA PCR Core Kit (Applied Biosystems) according to the manufacturer's instructions. PCR analysis was done using 25-μL volumes with SYBR Green PCR Core Reagents (Applied Biosystems). Primers (200 nmol/L) and RT mix (2 μL) were used in each PCR. Each sample was assayed in triplicate per PCR run and the experiment was repeated thrice. The cycling conditions were an initial denaturation at 95°C for 10 min, 40 cycles of denaturation at 95°C for 15 s, and elongation at 60°C for 60 s. The real-time PCR data were analyzed using the comparative C_t method.

Transfection of small interfering RNA against NF-κB p65 and transduction of lentiviral short hairpin RNA against CREB

SMARTpool-sequenced small interfering RNA (siRNA) targeting human NF-κB p65 (GenBank accession no. NM_021975) and nonspecific control pool siRNA were purchased from Dharmacon RNA Technologies and diluted to 20 μmol/L. NSCLC cells at 50% confluence were transfected with siRNA for NF-κB p65 or control siRNA at final concentrations of 50 and 100 nmol/L using the LipofectAMINE 2000 transfection reagent (Invitrogen) according to the manufacturer's instructions. Seventy-two hours after transfection, the cells were treated with IL-1β for 8 h. Total protein and RNA were collected from each sample.

For CREB-targeting viral short hairpin RNA (shRNA) delivery, the lentiviral plasmid pLKO.1 with a shRNA clone against CREB (clone ID: TRCN0000011085) was purchased from Open Biosystems. The pLKO.1 plasmid with a scrambled shRNA sequence and virus packaging plasmids (psPAX2 and pseudo-typing plasmid pMD2.G) were obtained from Addgene. HEK 293T cells were obtained from the American Type Culture Collection. HEK 293T cells at 50% confluence were cotransfected with pLKO.1 and virus packaging plasmids using FuGENE 6 transfection reagent (Roche Applied Science). After transfection for 16 h, the medium containing the transfection agent was replaced with fresh growth medium. Viral particles were then collected from the medium every 24 h twice. The virus titer in the pooled suspension was determined by counting the puromycin-resistant colonies in the virus-transduced culture. For knockdown of CREB, cells at 50% confluence were incubated with viral suspension at a multiplicity of infection of about 50 for 16 h. Four days after transduction, the cells were treated with IL-1β for 24 h and total protein and RNA were collected from the cells.

Migration assay

HUVECs (5×10^4) were suspended in serum-free RPMI 1640 and seeded in transwell chambers coated with gelatin. CM from NSCLC cells were applied in the outer chambers. HUVECs were incubated at 37°C in 5% CO_2 for 16 h. Following incubation, cells were fixed in 90% ethanol and stained with 0.1% crystal violet. Nonmigrated cells on the upper surface of the chamber filters were removed by swabbing, and the cells that had migrated through the filter were photographed under a microscope and quantified using the ImageJ software program (NIH, Bethesda, MD).[1]

Western blot analysis

Western blot analysis of target proteins was done as described previously (30). Equal amounts of protein (30 μg) were resolved using 10% SDS-PAGE. The mouse mAb against human NF-κB p65 was diluted in 5% nonfat milk at a ratio of 1:200 and rabbit polyclonal antibodies against human CREB and p-CREB were diluted at a ratio of 1:1,000 and incubated with the membranes overnight at 4°C. Proteins reactive with the primary antibody were visualized with a horseradish peroxidase–conjugated goat anti-mouse or goat anti-rabbit secondary antibody and enhanced chemiluminescence reagents (Amersham Bioscience).

Measurement of CXC chemokine protein secretion

The levels of secreted CXC chemokines in the CM were measured with immuno-dot blotting and ELISA. Immuno-dot blotting was done as described before (24, 28). Briefly, the CM were applied to a nitrocellulose membrane using the Manifold I Dot-Blot System (Scheleicher & Schuell). The membrane was then probed with antihuman CXCL5 and CXCL8 antibodies and the target proteins were detected by chemiluminescence and quantified by densitometry. ELISA was done with Quantikine CXCL5 and CXCL8 ELISA kits (R&D Systems) according to the manufacturer's instruction. Both measurements were normalized against the cell number.

Chromatin immunoprecipitation analysis

Chromatin immunoprecipitation assay was done using EZ ChIP kits (Upstate) according to the manufacturer's instruction. Briefly, NHTBE, BEASE-2B, 1799, 1198, and 1170-1 cells were grown in plate with normal media and chromatins were cross-linked by reaction with 1% formaldehyde for 10 min. The cross-linked chromatins were fragmented by sonication and subsequently immunoprecipitated with anti-CREB (Upstate) or anti–NF-κB p65 (Santa Cruz Biotechnology). The DNA in the precipitate was purified and used as the template for PCR. The primers for *CXCL5* promoter with NF-κB binding site were 5′-TAGAGGTGCACGCAGCTCCT (forward) and 5′-GAGCACTGTGGCTTCCTCGT (reverse); for *CXCL5* promoter with CREB binding site, 5′-CTGGACACACGTATACTTGC (forward) and 5′-GGCAGGTCATTCTAGGTTTC (reverse); and for *CXCL8* promoter with both CREB and NF-κB binding sites, 5′-AAAACTTTCGTCATACTCCG (forward) and 5′-AAAGTTTGTGCCTTATGGAG (reverse). PCR products were then separated in 1.2% agarose gel and stained with GelRed (Biotium).

Statistical analysis

Each experiment presented in the figures was repeated three or more times. The data are presented as the mean ± SE. Comparisons between groups were evaluated using ANOVA and a two-tailed Student's *t* test. $P < 0.05$ was considered statistically significant.

Results

Microarray analysis revealed an array of CXC chemokine genes up-regulated by IL-1β in NHTBE and A549 cells

To identify IL-1β–responsive genes involved in angiogenesis and tumorigenesis in NSCLC, we used total RNA isolated from NHTBE and A549 cells treated with or without 2.5 ng/mL IL-1β for 8 hours to generate cRNA for microarray hybridization and analysis using 44K whole human

---

[1] http://rsb.info.nih.gov/ij

Table 1. Analysis of IL-1β induction of CXC chemokine gene expression in NHTBE and A549 cells using microarray and real-time PCR analysis

Gene name	Systematic name	Gene description	NHTBE (fold change)		A549 (fold change)	
			Microarray	PCR	Microarray	PCR
CXCL1	NM_001511	*Homo sapiens* chemokine (C-X-C motif) ligand 1	18.8	11.9	32.5	26.9
CXCL2	NM_002089	*Homo sapiens* chemokine (C-X-C motif) ligand 2	18.8	12.9	23.3	14.9
CXCL3	NM_002090	*Homo sapiens* chemokine (C-X-C motif) ligand 3	18.2	29.9	19.2	11.7
CXCL5	NM_002994	*Homo sapiens* chemokine (C-X-C motif) ligand 5	173.3	95.9	12.0	8.2
CXCL6	NM_002993	*Homo sapiens* chemokine (C-X-C motif) ligand 6	31.3	80.6	13.9	28.8
CXCL8	NM_000584	*Homo sapiens* chemokine (C-X-C motif) ligand 8	11.9	17.6	40.8	32.4
VEGF	NM_001025366	*Homo sapiens* vascular endothelial growth factor	2.49	1.93	3.2	2.14

genome oligonucleotide microarrays. Analysis of the resulting microarray images using the Agilent feature extraction software program showed that the expression of six CXC chemokine genes, *CXCL1*, *CXCL2*, *CXCL3*, *CXCL5*, *CXCL6*, and *CXCL8*, was commonly up-regulated in response to IL-1β in both NHTBE and A549 cells ($P < 0.001$). As expected, we also saw that the expression of another angiogenic gene, *VEGF*, was up-regulated by IL-1β in both NHTBE and A549 cells (Table 1). To confirm that these microarray data reflect the IL-1β–induced gene expression, we examined the expression of *CXCL1*, *CXCL2*, *CXCL3*, *CXCL5*, *CXCL6*, and *CXCL8* genes in NHTBE and A549 cells using quantitative reverse transcription-PCR (RT-PCR). Indeed, the expression of all of these CXC chemokine genes was strongly induced after stimulation by IL-1β, and the quantitative RT-PCR results agreed with the microarray analysis results (Table 1).

IL-1β differentially regulates CXCL5 and CXCL8 protein secretion in NHTBE and NSCLC cells

According to the microarray and quantitative RT-PCR data, IL-1β stimulated the gene expression of angiogenic CXC chemokines in both NHTBE and A549 cells. We next verified whether such induction of gene expression was indeed translated into protein secretion in these cells. For this purpose, we focused on CXCL5 and CXCL8 because they reportedly have dominant angiogenic effects in lung cells. Using the immunodot blot assay, we detected the secretion of CXCL5 and CXCL8 in the CM from all four NSCLC cell lines (H226, H2170, A549, and H1734) treated with or without IL-1β. However, by the same method, the secretion of CXCL8 in NHTBE cells was detected only in IL-1β–treated cells, and not in untreated cells, and the CXCL5 protein was not detectable even with IL-1β treatment (data not shown). We verified such differential expression of angiogenic CXC chemokines between tumor and normal cells with ELISA (Fig. 1A and B). Again, we observed a significantly greater expression of CXCL5 and CXCL8 in tumor cells than in normal cells, especially with the induction of IL-1β. This result indicated that although IL-1β increases CXCL5 mRNA level in both normal epithelial cells and lung cancer cells, only the lung cancer cells produce a significant level of CXCL5 chemokine protein. Because A549 cells produced substantial amounts of both CXCL5 and CXCL8, we used A549 cells as a model system to further explore the effect of IL-1β on angiogenesis.

To further assess the response of *CXCL5* and *CXCL8* gene expression to IL-1β treatment, we measured the time and dose effect of IL-1β on A549 cells. The level of *CXCL5* and *CXCL8* transcription peaked at 4 hours after IL-1β stimulation (Fig. 2A), indicating that the cells responded to IL-1β by rapidly inducing the expression of CXC chemokine genes.

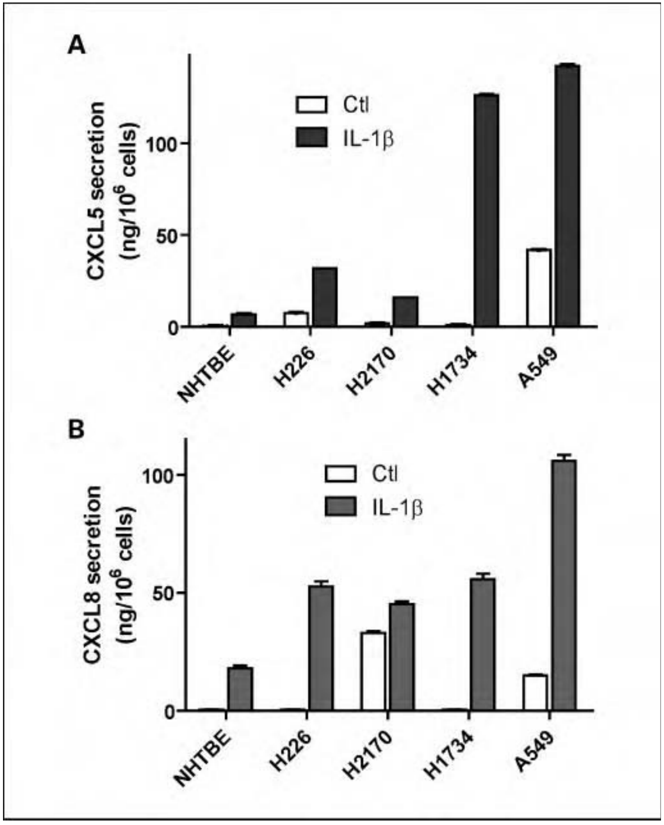


Fig. 1. IL-1β differentially regulates CXCL5 and CXCL8 protein secretion in NHTBE and NSCLC cells. NHTBE, H226, H2170, H1734, and A549 cells were treated with 2.5 ng/mL IL-1β for 24 h. CM were collected from these cell lines. The concentrations of CXCL5 (*A*) and CXCL8 (*B*) in CM were measured by ELISA. Results were normalized according to the cell number. *Columns*, mean; *bars*, SE.

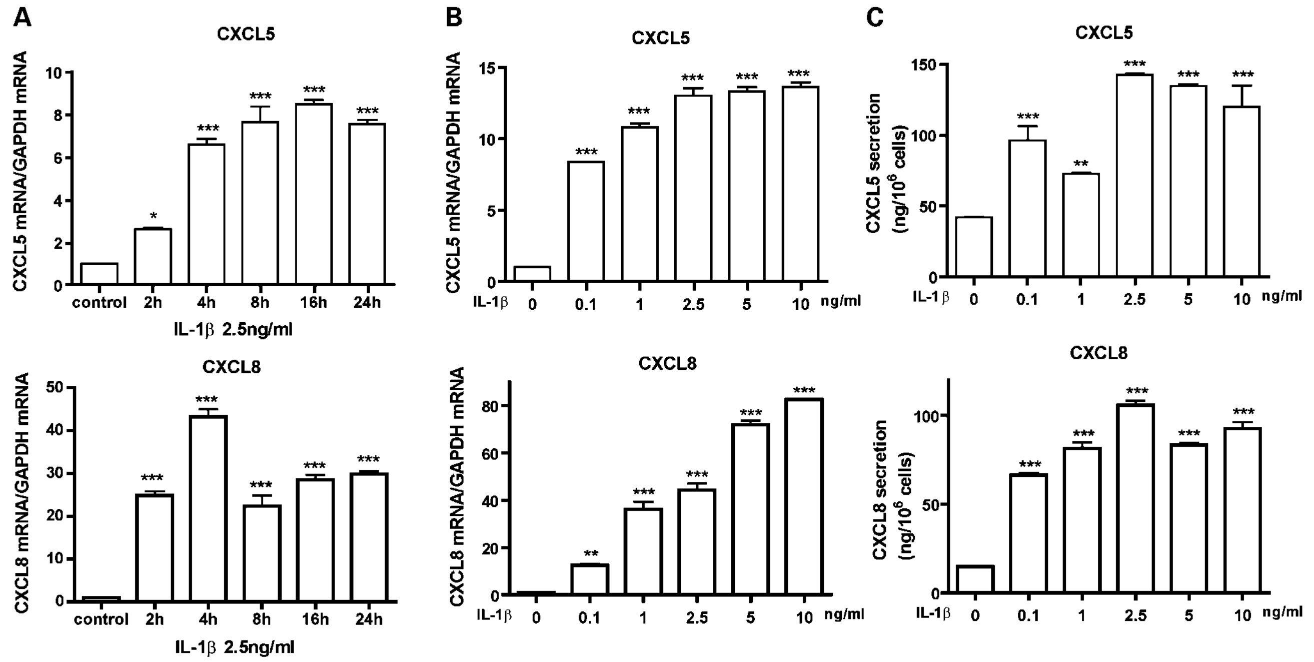


Fig. 2. IL-1β stimulates *CXCL5* and *CXCL8* gene expression and protein secretion in A549 cells in a time- and dose-dependent manner. *A*, A549 cells were incubated with control medium or with medium containing 2.5 ng/mL IL-1β for the indicated times. Total RNAs were collected from the cells and expression of the indicated genes was measured by quantitative RT-PCR. *GAPDH* was used as an internal control for normalizing the RNA loading. *B*, A549 cells were incubated with IL-1β at the indicated concentrations for 8 h. Total RNAs were collected from the cells for quantitative RT-PCR. *C*, culture media were collected from A549 cells with treatment as in *B* and the concentrations of CXCL5 and CXCL8 were measured with ELISA. *Columns*, mean; *bars*, SE. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, compared with untreated control.

We found that *CXCL5* mRNA expression was significantly induced even with IL-1β at 0.1 ng/mL. *CXCL8* mRNA expression reached a plateau at 2.5 to 5.0 ng/mL of IL-1β treatment (Fig. 2B). We confirmed this dose-dependent response by measuring the CXCL protein secretion with ELISA (Fig. 2C).

IL-1β significantly augments the angiogenic activity of NSCLC by inducing the expression of angiogenic CXC chemokine genes

Because IL-1β up-regulates the expression of important angiogenic CXC chemokine genes in NSCLC, we inferred that IL-1β may enhance the net angiogenic activity of NSCLC. To test this possibility, we conducted an endothelial cell (HUVEC) migration assay with CM from A549 NSCLC cell lines treated with or without IL-1β. As shown in Fig. 3A (*a* and *b*), compared with the CM from untreated A549 cells, the CM from IL-1β–treated cells (A549/IL-1β) markedly increased the angiogenic activity as assessed by the migration of HUVECs. Similar results were observed with CM prepared from IL-1β–treated H1734 cells (data not shown). These data clearly indicated that IL-1β significantly augments the net angiogenic activity of NSCLC as measured by the chemotactic activity of endothelial cell. Next, we hypothesized that the increased migration of HUVECs in response to CM from A549/IL-1β cells was attributable to the observed increased expression of angiogenic CXC chemokine genes in NSCLC cells induced by IL-1β. To test this, we conducted a HUVEC migration assay with CM from A549/IL-1β cells in the presence of neutralizing antibodies against CXCL5, CXCL8, CXCR2, and VEGF. We found that the migration of HUVECs was reduced significantly in the presence of neutralizing antibodies against CXCL5 and CXCR2 (Fig. 3A, *d* and *f*). However, we did not observe a similar effect on antibodies against CXCL8 and VEGF (Fig. 3A, *e* and *g*). These results suggested that the IL-1β–induced angiogenic activity of NSCLC is mainly attributable to the induced expression of CXCR2-dependent CXC chemokines, most likely CXCL5.

NF-κB and CREB mediate IL-1β–induced CXC chemokine gene expression

To investigate the mechanism underlying the IL-1β–induced CXC chemokine gene expression, we focused on the transcription factors mediating this effect. Sequence analysis of the promoters of these CXC chemokine genes indicated the potential binding sites for NF-κB, activator protein 1, activator protein 2, activator protein 3, Sp1, IFN regulatory factor 1, hepatic nuclear factor 1, and CREB (Fig. 4A; refs. 31–34). Because IL-1β activates NF-κB and CREB, and the NF-κB site and CRE-like sites are located in the promoters of *CXCL1, CXCL2, CXCL3, CXCL5,* and *CXCL8* genes, we sought to determine whether CREB and NF-κB mediate IL-1β–induced CXC chemokine gene expression in NSCLC cells. For this purpose, we abrogated NF-κB and CREB gene expression in A549 and H1734 cells by either transfecting the cells with a siRNA for NF-κB p65 or transducing the cells with a lentivirus containing the shRNA for CREB (shCREB). We confirmed that the NF-κB and CREB protein

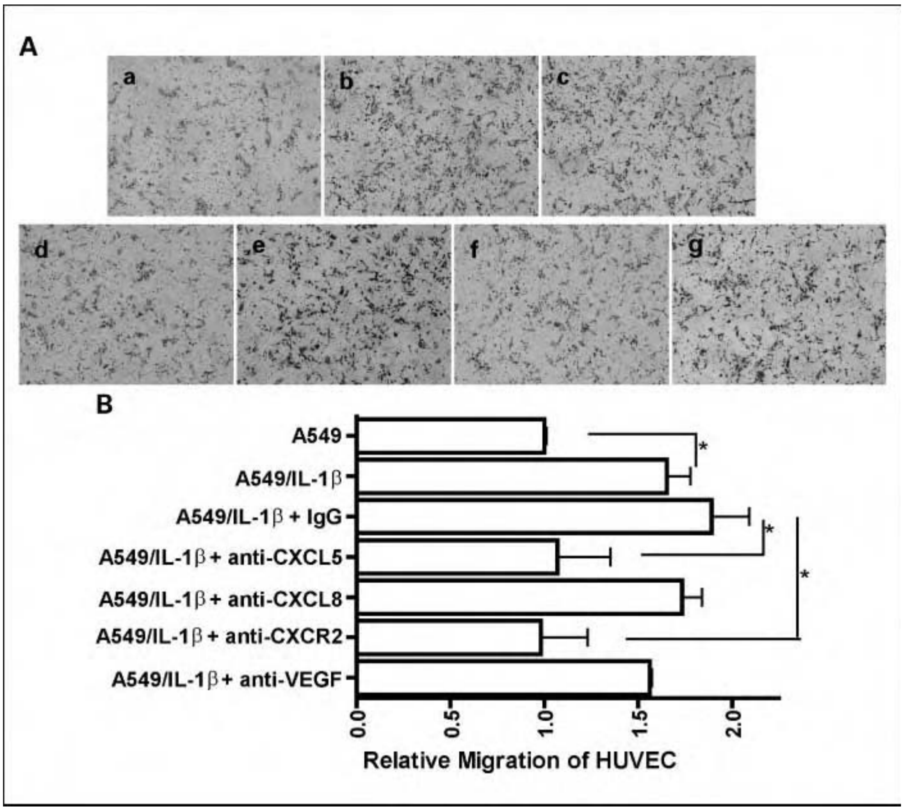


Fig. 3. IL-1β significantly augments the angiogenic activity of NSCLC by inducing the expression of angiogenic CXC chemokine genes. *A*, CM were collected from A549 cells treated with either control medium (*a*) or medium containing 2.5 ng/mL IL-1β (*b*) for 24 h, and CM from IL-1β–treated A549 cells were preincubated with 10 μg/mL of nonimmune IgG (*c*), an anti-CXCL5 mAb (*d*), an anti-CXCL8 mAb (*e*), an anti-CXCR2 mAb (*f*), or an anti-VEGF mAb (*g*) at 37°C for 1 h. A migration assay was done by stimulating HUVECs with the indicated CM as described in Materials and Methods. Migrated cells were photographed after 16 h of incubation at 37°C. *B*, image analysis of HUVEC migration was quantified using the ImageJ software program. *Columns*, mean; *bars*, SE. *, $P < 0.05$.

levels were knocked down by >80% after siRNA transfection or viral transduction in these two cell lines (Fig. 4B). We then treated the cells with IL-1β and measured the CXC chemokine gene expression levels in the cells. Real-time PCR results showed that IL-1β–induced CXC chemokine gene expression significantly decreased after knockdown of NF-κB (Fig. 4C). At the same time, shCREB-transduced cells had no or a much less significant response to IL-1β in the induction of CXC chemokine gene expression when compared with the nontransduced control or scrambled shRNA–transduced cells (Fig. 4D). These findings suggested that both CREB and NF-κB mediate IL-1β–induced CXC chemokine gene expression in NSCLC cells.

The expression level of CXC chemokine genes is associated with the development of lung cancer

To investigate if the angiogenic CXC chemokine genes are involved in the development of lung cancer, we measured the expression levels of CXCL5 and CXCL8 in an *in vitro* lung carcinogenesis model (IVLCM) that includes normal (NHTBE), immortalized (BEAS-2B and 1799), transformed (1198), and tumorigenic (1170-I) human bronchial epithelial cells. The quantitative RT-PCR results (Fig. 5A) showed that these two CXC chemokine genes were differentially expressed in the cell lines of IVLCM, with progressive increase from NHTBE to BEAS-2B and 1799 and then further to 1198 and 1170-I. As we showed that CREB and NF-κB regulate CXC chemokine gene expression, we examined whether the activities of these transcription factors correspond to the expression levels of CXCL5 and CXCL8 in these IVLCM cells. Western blot analysis and quantitative RT-PCR data showed that both the activity and the expression level of CREB were gradually increased with the progression of tumor development in the IVLCM (Fig. 5B and C). A similar pattern was also observed with NF-κB level (data not shown). Such trend of change was consistent with the tendency of their expression levels of CXCL5 and CXCL8. We further conducted a chromatin immunoprecipitation assay to determine whether CREB and NF-κB regulate the gene expression of these two chemokines in the IVLCM cells. The results of the chromatin immunoprecipitation assay (Fig. 5D) showed that CREB bound to the promoter regions of both *CXCL5* and *CXCL8* in all these cells, and that whereas NF-κB also bound to the *CXCL8* promoter in all these cells, its binding to the *CXCL5* promoter was only detected in the tumorigenic (1170-I) cells. These results indicated that the expression of the angiogenic CXC chemokine genes in these cells is well correlated with the progression of lung cancer from normal to invasive phenotype and might be further associated with their inherent activities of CREB and NF-κB.

KG-501 inhibits the endothelial cell migration induced by CM from IL-1β–treated NSCLC cells via the suppression of CXC chemokine gene expression in NSCLC cells

In looking for inhibitors that can block IL-1β–induced endothelial cell migration, we focused on the agents targeting the transcription factors NF-κB and CREB. It has been reported that KG-501 is a small molecule that binds to the transcription coactivator CREB-binding protein (CBP) and blocks the interaction of CBP with the active form of CREB, phospho-CREB (35). KG-501 can also inhibit NF-κB transcription activity because NF-κB also uses CBP as a cofactor to regulate gene expression (36). Based on our results that CREB and NF-κB could mediate IL-1β–induced CXC chemokine gene expression in NSCLC cells, we hypothesized that KG-501 could suppress the expression of these CXC chemokine genes and inhibit the endothelial cell migration induced by CM from NSCLC/IL-1β cells. As shown in Fig. 6A and B, the migration of HUVECs induced by CM from A549 cells treated with IL-1β plus 10 μmol/L of KG-501 was significantly lower than that induced by CM from A549 cells treated with IL-1β alone ($P < 0.05$). Next, we evaluated the effect of KG-501 on the transcriptional and proteomic levels of CXC chemokines induced by IL-1β. At 10 μmol/L, KG-501 suppressed the expression of all of the IL-1β–induced CXC chemokine genes except *CXCL8* (Fig. 6C). For the protein level, we measured the expression of CXCL5 and CXCL8 in A549 cells. KG-501 significantly suppressed IL-1β–induced CXCL5 protein secretion. However, its effect on IL-1β–induced CXCL8 protein secretion was not consistent with its concentration, with a stimulatory effect at low concentrations but a slightly inhibitory effect at high concentrations (Fig. 6D). Similar effects of KG-501 were also observed in the H1734 cell line (data not shown).

Discussion

In the present study, we showed that the proinflammatory cytokine IL-1β up-regulates the expression of an array of proangiogenic CXC chemokine genes in the NSCLC cell lines and that both of the transcription factors CREB and NF-κB can mediate this up-regulation. IL-1β augments the angiogenic activity of NSCLC, as manifested by the ability of CM from IL-1β–treated cells to induce endothelial cell migration. Our finding that the transcription factors CREB and NF-κB mediate IL-1β–induced CXC chomokine gene expression extends our knowledge about the mechanism of gene regulation by the angiogenic factors and provides new potential targets for angioprevention. Finally, the findings that the small molecule KG-501 significantly suppressed IL-1β–induced CXC chemokine gene expression and in turn reduced the CM-induced endothelial cell migration indicate that KG-501 may have therapeutic and preventive potential for NSCLC.

Our data showing that IL-1β up-regulates the expression of angiogenic CXC chemokine genes and augments the angiogenic activity of NSCLC are consistent with previous reports on the role of CXCR2 and its ligands in promoting tumor-associated angiogenesis and early development of NSCLC (22, 23, 37–40). In an *in vivo* study using murine Lewis lung cancer heterotopic and orthotopic tumor model systems with CXCR2$^{-/-}$ versus CXCR2$^{+/+}$ mice, researchers showed that the tumors in CXCR2$^{-/-}$ mice exhibited reduced growth, increased necrosis, inhibited tumor-associated angiogenesis, and reduced metastatic potential (37). Similar to our finding that a neutralizing antibody against CXCR2 blocked CM-induced endothelial cell migration, the report showed that a specific neutralizing antibody against CXCR2 inhibited tumor growth, increased necrosis, and reduced tumors vessel density in CXCR2$^{+/+}$ mice (37). Furthermore, studies showed that CXCL5 and CXCL8 play a dominant role in promoting angiogenesis in patients with NSCLC (22, 23, 40). Whereas CXCL8 was the first angiogenic ELR-positive CXC chemokine discovered in NSCLC, CXCL5 reportedly has a higher degree of correlation with NSCLC-derived angiogenesis (23). In our study using the neutralizing antibodies, we observed that CXCL5 neutralization inhibited the migration of endothelial cells to the same degree as did CXCR2 neutralization. We failed to see this inhibitory effect using the CXCL8-neutralizing antibody, indicating that CXCL8 produced by A549 cells in response to IL-1β may not be sufficient to induce endothelial cell migration. Our experiment using KG-501 further supported such observation, as this small molecule blocked endothelial cell migration without affecting the CXCL8 level in CM. In addition, another angiogenic factor, VEGF, may play only a minor role in inducing endothelial cell migration in NSCLC, as neutralization of which could not inhibit the migration of HUVECs. Because high levels of CXCL5 and CXCL8 protein expression were detected in IL-1β–treated NSCLC cells but only a very low level of CXCL5 protein was induced in NHTBE cells, we speculated that IL-1β may induce an angiogenic response only in NSCLC tumor cells but not in surrounding normal cells.

It is well documented that IL-1β up-regulates the expression of the proangiogenic CXC chemokine genes and that NF-κB is the common transcription factor that mediates this effect (31, 41, 42). All of the angiogenic CXC chemokine gene promoters contain a putative *cis*-element that is recognized by the NF-κB family of transcriptional factors (31, 33, 43). Consistent with these findings, our results showed that after knockdown of NF-κB p65 by siRNA transfection, both basal and IL-1β–induced expression of CXC chemokine genes decreased dramatically in A549 and H1734 cells. IL-1β can also regulate gene expression through the transcription factor CREB. Previously, we reported that IL-1β activates the mitogen-activated protein kinase (ERK1/2)/mitogen- and stress-activated protein kinase/CREB pathway and regulates *MUC5AC* gene expression in human airway epithelial cells (30). Sequence analysis of the CXC chemokine gene promoters identified a CRE or CRE-like domain in the gene *CXCL1, CXCL2, CXCL3, CXCL5, and CXCL8* (31, 32, 34). In the present study, we showed that NSCLC cells with CREB knockdown were much less responsive to IL-1β than cells without such knockdown in terms of the induction of angiogenic CXC chemokine gene expression. We further confirmed the binding of CREB to the *CXCL5* and *CXCL8* promoters with chromatin immunoprecipitation assay. These data suggested that CREB can also regulate the expression of CXC chemokine genes. It has been reported that cyclooxygenase-2 is critical for IL-1β–induced angiogenesis both *in vitro* and *in vivo* through the production of prostanoids such as prostaglandin E_2 and thromboxane A_2 (44). In addition, Pold et al. (40) reported that cyclooxygenase-2 contributes to the progression of NSCLC tumorigenesis by enhancing the expression of *CXCL5* and *CXCL8*. Because cyclooxygenase-2 expression is also regulated by CREB in many cell types, including lung cell lines (45–48), CREB may regulate CXC chemokine gene expression through multiple mechanisms that need to be further delineated. In addition, the interaction between CREB and NF-κB in regulating the expression of these chemokine genes should be further investigated because these two transcription factors may compete with each other for the binding to CBP or work synergistically to affect the outcome of gene regulation (49). Additionally, the response elements for CREB and NF-κB on the promoter of *CXCL8* gene are consecutively

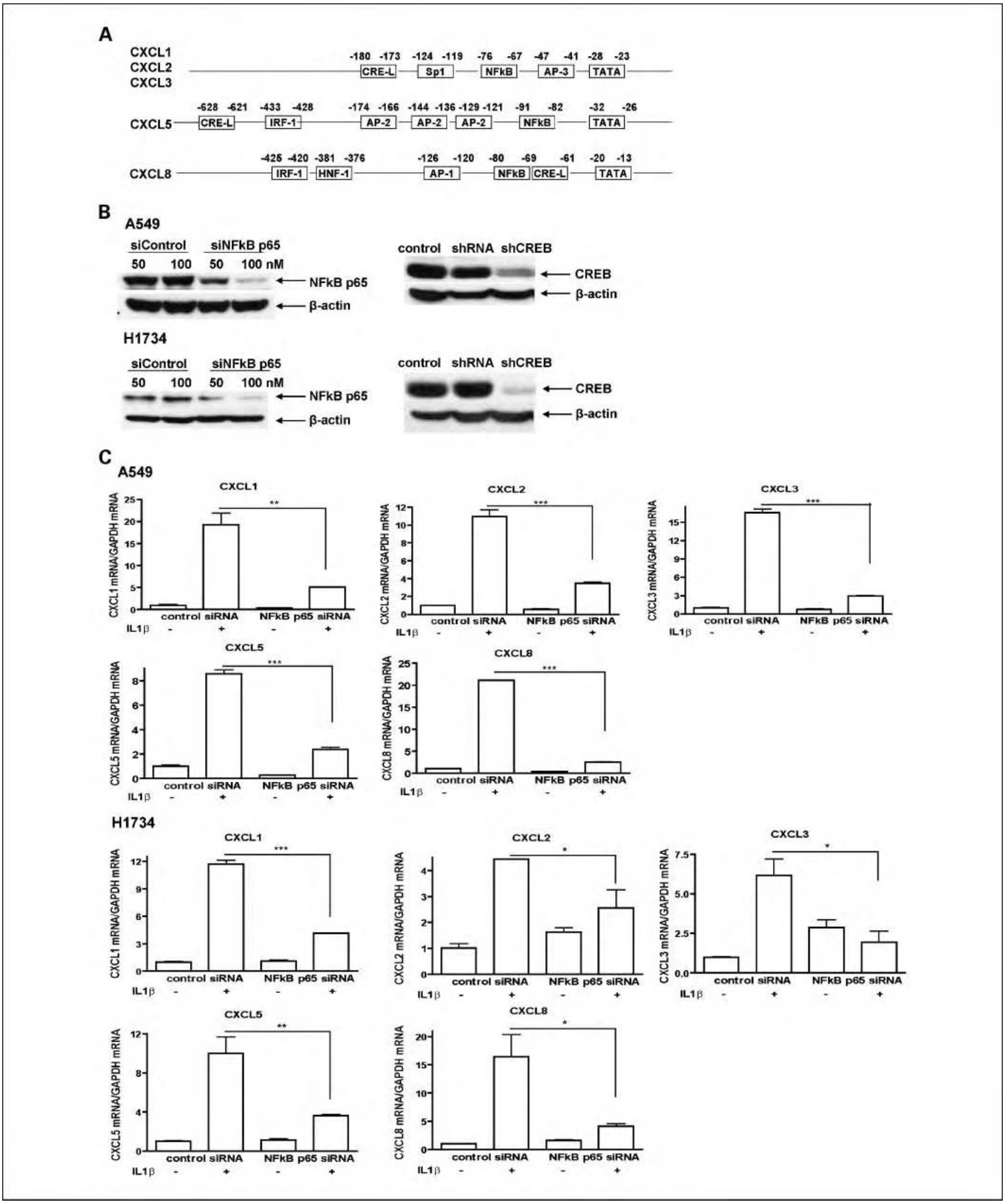


Fig. 4. NF-κB and CREB mediate IL-1β–induced CXC chemokine gene expression. *A,* transcription factor binding domains located in the gene promoters of *CXCL1, CXCL2, CXCL3, CXCL5,* and *CXCL8*. *B,* A549 cells were transfected with siRNA for NF-κB p65 or transduced with shCREB lentivirus; nontransduced cells and cells transfected with control siRNA or transduced with scrambled shRNA were used as controls. Proteins were extracted from the cells 72 h after transfection or 96 h after transduction. NF-κB and CREB protein expression levels were detected by Western blot analysis. β-Actin protein was probed as a loading control. *C,* A549 and H1734 cells were transfected with siRNA for NF-κB p65, and cells transfected with control siRNA were used as control. Seventy-two hours later, the cells were treated with either control medium or medium containing 2.5 ng/mL IL-1β for 8 h.

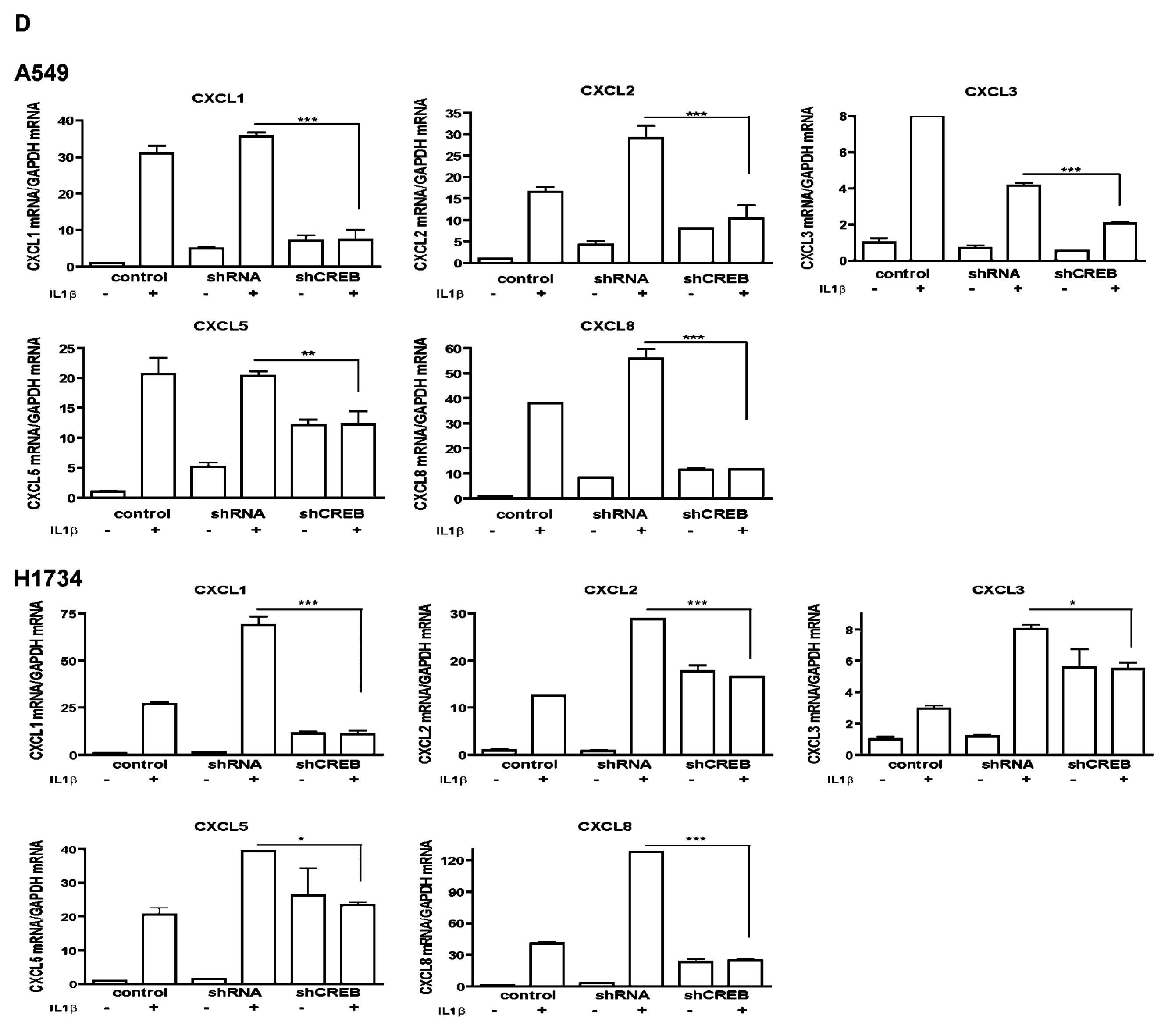


Fig. 4 *Continued. D*, A549 and H1734 cells were transduced with shCREB, and nontransduced cells and cells transduced with scrambled shRNA were used as controls. Ninety-six hours later, the cells were treated with or without 2.5 ng/mL IL-1β for 24 h. Total RNAs were isolated from the cells in *C* and *D*. Real-time PCR was done with specific primers for the indicated genes. *Columns*, mean; *bars*, SE. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

located, which further complicated the interaction of these two transcription factors and might have also contributed to the perplexing pattern of CXCL8 expression in response to KG-501 treatment.

In the IVLCM cell lines, we detected that the expression of the angiogenic CXC chemokine genes *CXCL5* and *CXCL8* increased progressively from normal (NHTBE), immortalized (BEAS-2B and 1799), and transformed (1198) to tumorigenic (1170-I) human bronchial epithelial cells. This is not unexpected because the angiogenic CXC chemokines CXCL5 and CXCL8 have been reported to be elevated in NSCLC tissues and their expression is related to tumor progression (23, 50, 51). Recently, accumulating evidence shows that CXC chemokines induce tumorigenesis by stimulating cell proliferation, mediating cell survival, promoting angiogenesis, and facilitating tumor cell migration and invasion (13, 52). In addition, we detected that the expression and activation of CREB increased correspondingly with the tumorigenicity in these IVLCM cell lines. The consistency of the expression and activation of these factors with the expression of angiogenic CXC chemokine genes in IVLCM confirmed the regulation of CXC chemokine genes by CREB and NF-κB, and lent support to their involvement in lung carcinogenesis. CREB has been known for its role in cell proliferation and survival (53–55). We have recently reported that the basal activity and expression level of CREB are commonly higher in a number of NSCLC cell lines versus normal cells, and the inhibition of CREB transcription activity induces apoptosis in these NSCLC cells by suppressing the expression of CREB-regulated genes that are involved in cell proliferation (56). Moreover, we have accumulated data from archived tumor tissue specimens showing that the CREB and p-CREB levels are commonly higher in the lung tumor tissues versus the adjacent normal tissues (57). By showing the role of CREB in tumor angiogenesis, the current study further suggests that CREB can be an effectual target for therapy as well as prevention of NSCLC.

KG-501 is a small molecule that binds to the KIX domain of CBP (35). It disrupts the interaction of CBP with CREB and inhibits the CREB-dependent activation of cellular genes. KG-501 can also disrupt the interaction of other factors with CBP, such as NF-κB (35). Our finding that KG-501 reduced the NSCLC/IL-1β conditioned medium–induced migration of HUVECs by down-regulating the expression of the ELR-positive CXC chemokine genes in NSCLC cells indicates that KG-501 can be used as a therapeutic and/or preventive agent for inhibiting tumor-associated angiogenesis in NSCLC. Although the *CXCL8* promoter contains both a NF-κB-binding site and a CRE-like motif, and our knockdown experiments show that the depletion of either factor reduced IL-1β–induced CXCL8 expression (Fig. 4C and D), KG-501 was not able to effectively inhibit *CXCL8* gene expression and protein secretion. Because KG-501 disrupts CREB-CBP or NF-κB-CBP interaction, the regulation of *CXCL8* gene by these factors may be mediated via a CBP-independent mechanism. A recent study showed that the transcriptional activity of CREB on *CXCL8* promoter requires a different coactivator, termed transducer of regulated CREB (TORC1; ref. 34), suggesting a different regulatory mechanism beyond the binding of CREB to CRE. Nevertheless, CXCL8 did not seem to play a critical role in the induction of HUVEC migration as evidenced by the results of CXCL8 neutralization; therefore, KG-501 can still effectively inhibit NSCLC/IL-1β CM–induced HUVEC migration.

Our findings also implicated a positive association between chronic obstructive pulmonary disease (COPD) and lung cancer. Chronic obstructive pulmonary disease is a product of chronic inflammation that leads to tissue damage and physiologic adaptations (58, 59). It has been known for years that local inflammation in the lungs plays an important role in airway remodeling and parenchymal destruction, which are effects typified by chronic obstructive pulmonary disease (59). It is now well recognized that, in addition to lung inflammation, patients with chronic obstructive pulmonary disease frequently show persistent low-grade systemic inflammation, with the characteristic release of proinflammatory mediators such as IL-1β and tumor necrosis factor α into the circulation (60). Considerable evidence has associated chronic inflammation with cancer development. Our finding that CREB and NF-κB regulate proangiogenic CXC chemokines in response to proinflammatory cytokine IL-1β may provide a novel mechanistic linkage between chronic obstructive pulmonary disease and the development of lung cancer.

In summary, IL-1β increases the angiogenic activity of NSCLC by up-regulating the expression of an array of proangiogenic CXC chemokine genes, which subsequently induces endothelial cell migration. The transcription factors CREB and NF-κB both can mediate this effect, suggesting that these two transcription factors are involved in tumor-associated angiogenesis and, therefore, could be potential targets for the angioprevention in NSCLC. We also conclude that the

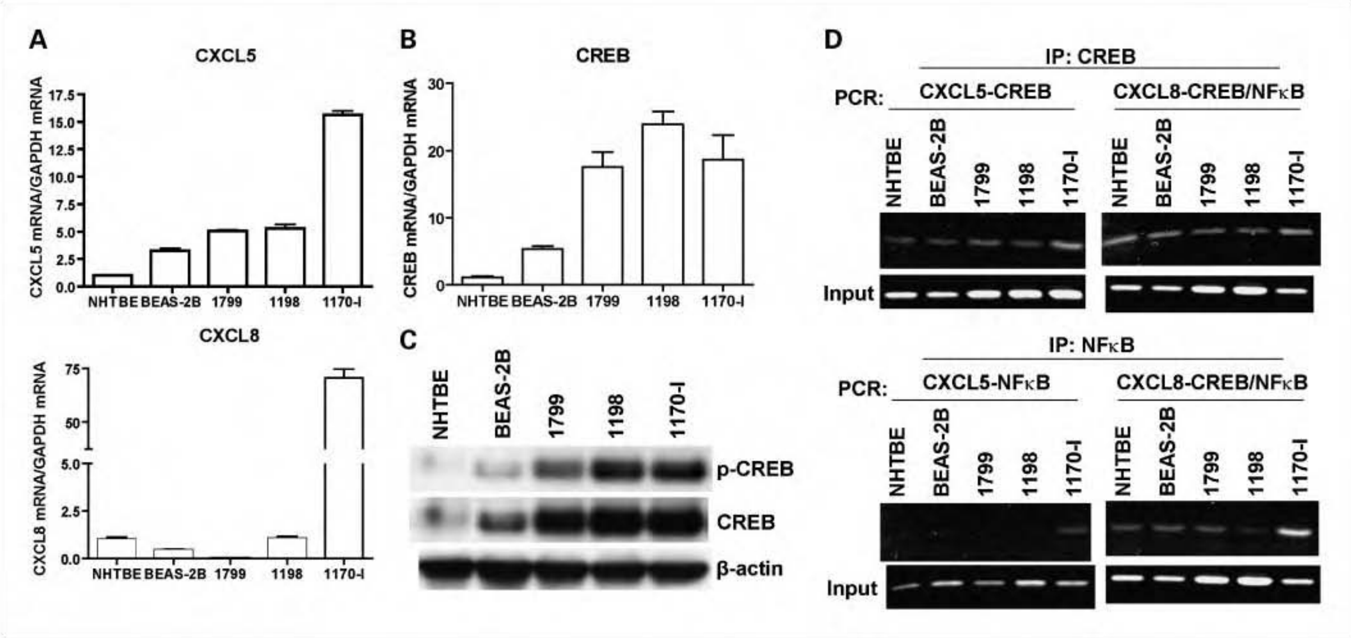


Fig. 5. The expression of CXC chemokine genes and CREB is associated with the development of lung cancer. *A* and *B*, total RNAs from NHTBE and subconfluent cultures of BEAS-2B, 1799, 1198, and 1170-I were extracted and the expressions of designated genes were quantitated with real-time PCR using specific primers. Each specific gene expression was normalized with the expression level of *GAPDH. Columns*, mean; *bars*, SE. *C,* whole-cell lysates (20 μg/lane) were isolated from NHTBE and IVLCM cell lines, and the levels of phospho-CREB (p-CREB) and total CREB were measured by immunoblot analysis. *D*, CREB and NF-κB binding on CXCL5 and CXCL8 promoter. Chromatin immunoprecipitation assay was done with chromatins prepared from NHTBE, BEASE-2B, 1799, 1198, and 1170-1 cells. The binding of CREB or NF-κB to the CXCL promoter was detected by visualization of the PCR product. The single bands detected in input samples indicate the specificity of the PCR primers.

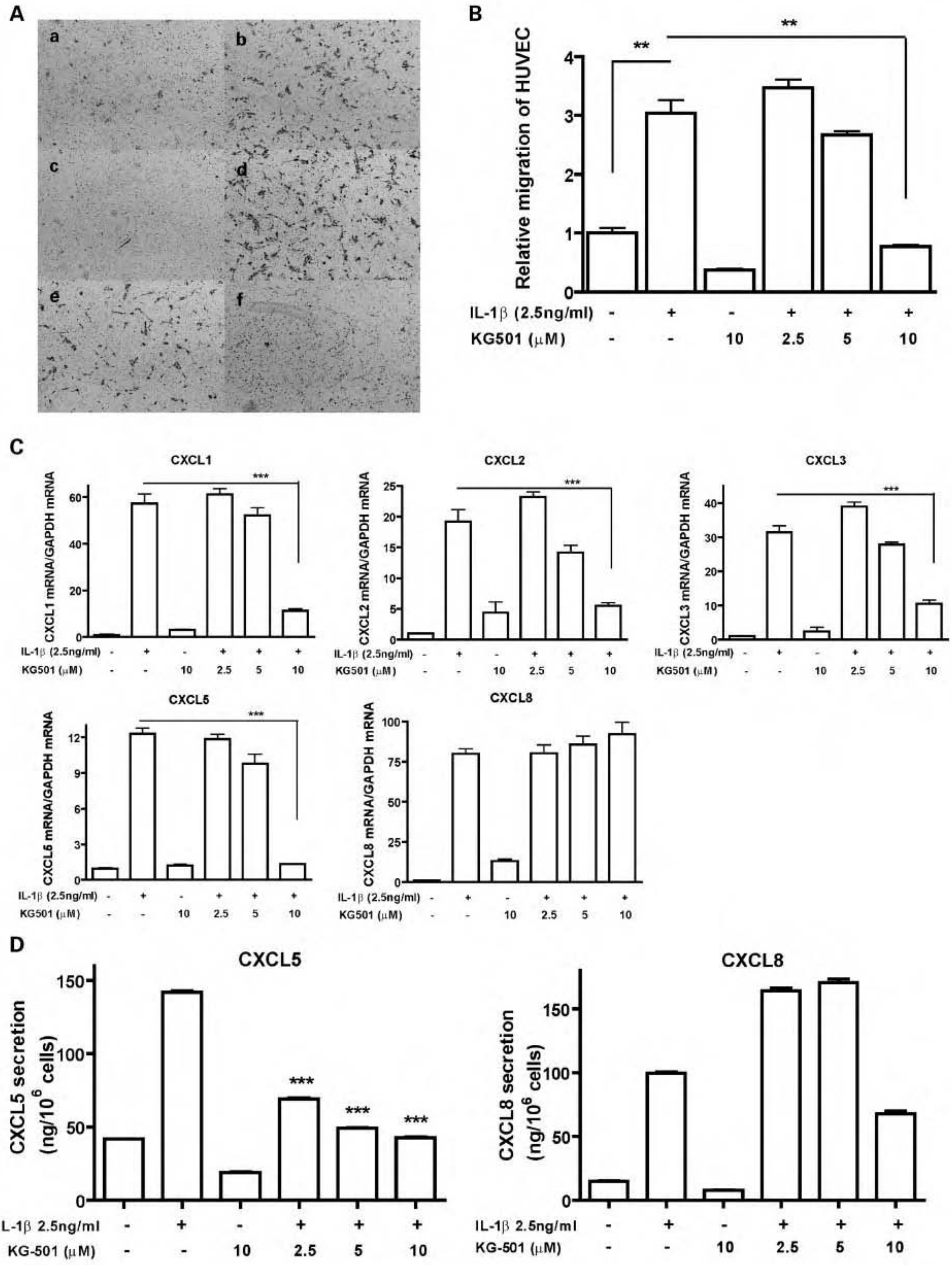


Fig. 6. KG-501 suppresses NSCLC/IL-1β CM–induced migration of HUVECs by regulating IL-1β–induced CXC chemokine gene expression in NSCLC cells. *A,* HUVEC migration was induced with CM from A549 cells treated with control medium (*a*) or with medium containing 2.5 ng/mL IL-1β (*b*), 10 μmol/L KG-501 (*c*), 2.5 ng/mL IL-1β plus 2.5 μmol/L KG-501 (*d*), 2.5 ng/mL IL-1β plus 5 μmol/L KG-501 (*e*), or 2.5 ng/mL IL-1β plus 10 μmol/L KG-501 (*f*) in serum-free medium for 24 h. *B,* quantification of image analysis for the HUVEC migration assay. *C,* total RNAs were isolated from cells with the treatment described in *A*, and real-time PCR was done to measure the expression level of the indicated genes. *D,* the concentration of CXCL5 and CXCL8 protein in the CM was measured with ELISA. Results were normalized according to the cell number. *Columns*, mean; *bars*, SE. **, $P < 0.01$; ***, $P < 0.001$.

small molecule KG-501 neutralizes the effect of NSCLC/IL-1β–conditioned medium on endothelial cell migration by inhibiting CREB and NF-κB transcriptional activity, which results in the down-regulation of CXC chemokine gene expression in NSCLC cells, suggesting that KG-501 may be used as a therapeutic and angiopreventive agent for NSCLC.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Don Norwood for critical editing of the manuscript.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008;58:71–96.
2. Folkman J. What is the evidence that tumors are angiogenesis dependent? J Natl Cancer Inst 1990; 82:4–6.
3. Hanahan D, Folkman J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell 1996;86:353–64.
4. Liotta LA, Stracke ML. Tumor invasion and metastases: biochemical mechanisms. Cancer Treat Res 1988;40:223–38.
5. Holmgren L, O'Reilly MS, Folkman J. Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression. Nat Med 1995;1:149–53.
6. Folkman J. Tumor angiogenesis: therapeutic implications. N Engl J Med 1971;285:1182–6.
7. Tosetti F, Ferrari N, De Flora S, Albini A. Angioprevention: angiogenesis is a common and key target for cancer chemopreventive agents. FASEB J 2002; 16:2–14.
8. Bisacchi D, Benelli R, Vanzetto C, Ferrari N, Tosetti F, Albini A. Anti-angiogenesis and angioprevention: mechanisms, problems and perspectives. Cancer Detect Prev 2003;27:229–38.
9. Albini A, Tosetti F, Benelli R, Noonan DM. Tumor inflammatory angiogenesis and its chemoprevention. Cancer Res 2005;65:10637–41.
10. Yoneda J, Kuniyasu H, Crispens MA, Price JE, Bucana CD, Fidler IJ. Expression of angiogenesis-related genes and progression of human ovarian carcinomas in nude mice. J Natl Cancer Inst 1998;90:447–54.
11. Herbst RS, Fidler IJ. Angiogenesis and lung cancer: potential for therapy. Clin Cancer Res 2000;6: 4604–6.
12. Bremnes RM, Camps C, Sirera R. Angiogenesis in non-small cell lung cancer: the prognostic impact of neoangiogenesis and the cytokines VEGF and bFGF in tumours and blood. Lung Cancer 2006; 51:143–58.
13. Strieter RM, Belperio JA, Burdick MD, Sharma S, Dubinett SM, Keane MP. CXC chemokines: angiogenesis, immunoangiostasis, and metastases in lung cancer. Ann N Y Acad Sci 2004;1028:351–60.
14. Voronov E, Shouval DS, Krelin Y, et al. IL-1 is required for tumor invasiveness and angiogenesis. Proc Natl Acad Sci U S A 2003;100:2645–50.
15. Lin EY, Pollard JW. Role of infiltrated leucocytes in tumour growth and spread. Br J Cancer 2004;90: 2053–8.
16. Brigati C, Noonan DM, Albini A, Benelli R. Tumors and inflammatory infiltrates: friends or foes? Clin Exp Metastasis 2002;19:247–58.
17. Bingle L, Brown NJ, Lewis CE. The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. J Pathol 2002;196:254–65.
18. Dirkx AE, Oude Egbrink MG, Wagstaff J, Griffioen AW. Monocyte/macrophage infiltration in tumors: modulators of angiogenesis. J Leukoc Biol 2006; 80:1183–96.
19. Nakao S, Kuwano T, Tsutsumi-Miyahara C, et al. Infiltration of COX-2-expressing macrophages is a prerequisite for IL-1 β-induced neovascularization and tumor growth. J Clin Invest 2005;115: 2979–91 Epub 005 Oct 20.
20. Dinarello CA. Biologic basis for interleukin-1 in disease. Blood 1996;87:2095–147.
21. Saijo Y, Tanaka M, Miki M, et al. Proinflammatory cytokine IL-1 β promotes tumor growth of Lewis lung carcinoma by induction of angiogenic factors: *in vivo* analysis of tumor-stromal interaction. J Immunol 2002;169:469–75.
22. Arenberg DA, Kunkel SL, Polverini PJ, Glass M, Burdick MD, Strieter RM. Inhibition of interleukin-8 reduces tumorigenesis of human non-small cell lung cancer in SCID mice. J Clin Invest 1996;97: 2792–802.
23. Arenberg DA, Keane MP, DiGiovine B, et al. Epithelial-neutrophil activating peptide (ENA-78) is an important angiogenic factor in non-small cell lung cancer. J Clin Invest 1998;102:465–72.
24. Gray TE, Guzman K, Davis CW, Abdullah LH, Nettesheim P. Mucociliary differentiation of serially passaged normal human tracheobronchial epithelial cells. Am J Respir Cell Mol Biol 1996; 14:104–12.
25. Kim SW, Cheon K, Kim CH, et al. Proteomics-based identification of proteins secreted in apical surface fluid of squamous metaplastic human tracheobronchial epithelial cells cultured by three-dimensional organotypic air-liquid interface method. Cancer Res 2007;67:6565–73.
26. Kolodziejski PJ, Musial A, Koo JS, Eissa NT. Ubiquitination of inducible nitric oxide synthase is required for its degradation. Proc Natl Acad Sci U S A 2002;99:12315–20.
27. Koo JS, Jetten AM, Belloni P, Yoon JH, Kim YD, Nettesheim P. Role of retinoid receptors in the regulation of mucin gene expression by retinoic acid in human tracheobronchial epithelial cells. Biochem J 1999;338:351–7.
28. Koo JS, Yoon JH, Gray T, Norford D, Jetten AM, Nettesheim P. Restoration of the mucous phenotype by retinoic acid in retinoid-deficient human bronchial cell cultures: changes in mucin gene expression. Am J Respir Cell Mol Biol 1999;20:43–52.
29. Lacroix L, Feng G, Lotan R. Identification of genes expressed differentially in an *in vitro* human lung carcinogenesis model. Cancer Biol Ther 2006; 5:665–73.
30. Song KS, Lee WJ, Chung KC, et al. Interleukin-1 β and tumor necrosis factor-α induce MUC5AC overexpression through a mechanism involving ERK/p38 mitogen-activated protein kinases-MSK1-CREB activation in human airway epithelial cells. J Biol Chem 2003;278:23243–50.
31. Anisowicz A, Messineo M, Lee SW, Sager R. An NF-κB-like transcription factor mediates IL-1/TNF-α induction of gro in human fibroblasts. J Immunol 1991;147:520–7.
32. Corbett MS, Schmitt I, Riess O, Walz A. Characterization of the gene for human neutrophil-activating peptide 78 (ENA-78). Biochem Biophys Res Commun 1994;205:612–7.
33. Mukaida N, Shiroo M, Matsushima K. Genomic structure of the human monocyte-derived neutrophil chemotactic factor IL-8. J Immunol 1989;143: 1366–71.
34. Iourgenko V, Zhang W, Mickanin C, et al. Identification of a family of cAMP response element-binding protein coactivators by genome-scale functional analysis in mammalian cells. Proc Natl Acad Sci U S A 2003;100:12147–52.
35. Best JL, Amezcua CA, Mayr B, et al. Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. Proc Natl Acad Sci U S A 2004;101:17622–7.
36. Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, Collins T. CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A 1997;94:2927–32.
37. Keane MP, Belperio JA, Xue YY, Burdick MD, Strieter RM. Depletion of CXCR2 inhibits tumor growth and angiogenesis in a murine model of lung cancer. J Immunol 2004;172:2853–60.
38. Wislez M, Fujimoto N, Izzo JG, et al. High expression of ligands for chemokine receptor CXCR2 in alveolar epithelial neoplasia induced by oncogenic kras. Cancer Res 2006;66:4198–207.
39. Addison CL, Daniel TO, Burdick MD, et al. The CXC chemokine receptor 2, CXCR2, is the putative receptor for ELR+ CXC chemokine-induced angiogenic activity. J Immunol 2000;165:5269–77.
40. Pold M, Zhu LX, Sharma S, et al. Cyclooxygenase-2-dependent expression of angiogenic CXC chemokines ENA-78/CXC ligand (CXCL) 5 and interleukin-8/CXCL8 in human non-small cell lung cancer. Cancer Res 2004;64:1853–60.
41. Keates AC, Keates S, Kwon JH, et al. ZBP-89, Sp1, and nuclear factor-κB regulate epithelial neutrophil-activating peptide-78 gene expression in Caco-2 human colonic epithelial cells. J Biol Chem 2001;276:43713–22.
42. Mukaida N, Mahe Y, Matsushima K. Cooperative interaction of nuclear factor-κB- and *cis*-regulatory enhancer binding protein-like factor binding elements in activating the interleukin-8 gene by pro-inflammatory cytokines. J Biol Chem 1990; 265:21128–33.
43. Rovai LE, Herschman HR, Smith JB. Cloning and characterization of the human granulocyte chemotactic protein-2 gene. J Immunol 1997;158: 5257–66.
44. Kuwano T, Nakao S, Yamamoto H, et al. Cyclooxygenase 2 is a key enzyme for inflammatory cytokine-induced angiogenesis. FASEB J 2004; 18:300–10.
45. Wardlaw SA, Zhang N, Belinsky SA. Transcriptional regulation of basal cyclooxygenase-2 expression in murine lung tumor-derived cell lines by CCAAT/enhancer-binding protein and activating transcription factor/cAMP response element-binding protein. Mol Pharmacol 2002;62:326–33.
46. Tang Q, Chen W, Gonzales MS, Finch J, Inoue H, Bowden GT. Role of cyclic AMP responsive element in the UVB induction of cyclooxygenase-2 transcription in human keratinocytes. Oncogene 2001;20:5164–72.
47. Han S, Sidell N, Roser-Page S, Roman J. Fibronectin stimulates human lung carcinoma cell growth by inducing cyclooxygenase-2 (COX-2) expression. Int J Cancer 2004;111:322–31.
48. Chen JJ, Huang WC, Chen CC. Transcriptional regulation of cyclooxygenase-2 in response to proteasome inhibitors involves reactive oxygen species-mediated signaling pathway and recruitment of CCAAT/enhancer-binding protein δ and CREB-binding protein. Mol Biol Cell 2005;16:5579–91.
49. Shenkar R, Yum HK, Arcaroli J, Kupfner J, Abraham E. Interactions between CBP, NF-κB, and CREB in the lungs after hemorrhage and

endotoxemia. Am J Physiol Lung Cell Mol Physiol 2001;281:L418–26.

50. Yuan A, Yang PC, Yu CJ, et al. Interleukin-8 messenger ribonucleic acid expression correlates with tumor progression, tumor angiogenesis, patient survival, and timing of relapse in non-small-cell lung cancer. Am J Respir Crit Care Med 2000;162:1957–63.

51. White ES, Flaherty KR, Carskadon S, et al. Macrophage migration inhibitory factor and CXC chemokine expression in non-small cell lung cancer: role in angiogenesis and prognosis. Clin Cancer Res 2003;9:853–60.

52. Yeudall WA, Miyazaki H. Chemokines and squamous cancer of the head and neck: targets for therapeutic intervention? Expert Rev Anticancer Ther 2007;7:351–60.

53. Conkright MD, Montminy M. CREB: the unindicted cancer co-conspirator. Trends Cell Biol 2005;15:457–9.

54. Shankar DB, Sakamoto KM. The role of cyclic-AMP binding protein (CREB) in leukemia cell proliferation and acute leukemias. Leuk Lymphoma 2004;45:265–70.

55. Abramovitch R, Tavor E, Jacob-Hirsch J, et al. A pivotal role of cyclic AMP-responsive element binding protein in tumor progression. Cancer Res 2004; 64:1338–46.

56. Aggarwal S, Kim SW, Ryu SH, Chung WC, Koo JS. Growth suppression of lung cancer cells by targeting cyclic AMP response element-binding protein. Cancer Res 2008;68:981–8.

57. Seo H-S, Liu D, Bekele B, et al. CREB overexpression: a feature associated with negative prognosis in never-smokers with NSCLC. Cancer Res In press 2008.

58. Szilasi M, Dolinay T, Nemes Z, Strausz J. Pathology of chronic obstructive pulmonary disease. Pathol Oncol Res 2006;12:52–60.

59. Jeffery PK. Remodeling and inflammation of bronchi in asthma and chronic obstructive pulmonary disease. Proc Am Thorac Soc 2004;1:176–83.

60. Hegab AE, Sakamoto T, Saitoh W, et al. Polymorphisms of TNFα, IL1β, and IL1RN genes in chronic obstructive pulmonary disease. Biochem Biophys Res Commun 2005;329:1246–52.

Cancer **Prevention** Research

Epidermal Growth Factor Receptor Abnormalities in the Pathogenesis and Progression of Lung Adenocarcinomas

Ximing Tang,[1] Marileila Varella-Garcia,[3] Ana C. Xavier,[3] Erminia Massarelli,[1] Natalie Ozburn,[1]
Cesar Moran[2] and Ignacio I. Wistuba[1,2]

Abstract To identify the characteristics and sequence of epidermal growth factor receptor (EGFR) abnormalities relevant to the pathogenesis and progression of lung adenocarcinoma, we performed a precise mapping analysis of *EGFR* mutation, gene copy number, and total and phosphorylated EGFR protein expression for the same tissue sites. We examined normal bronchial and bronchiolar epithelium (NBE) and tumor tissues obtained from 50 formalin-fixed lung adenocarcinomas, including 24 *EGFR*-mutant primary tumors with nine corresponding lymph node metastases and 26 wild-type primary tumors. NBE in 12 of 24 (50%) mutant and 3 of 26 (12%) wild-type tumors harbored *EGFR* mutations; these NBE also showed a lack of *EGFR* copy number increase and frequent EGFR (69%) and phosphorylated EGFR (33%) overexpression. *EGFR* mutation and protein overexpression were more frequent in NBE sites within tumors than in NBE sites adjacent to and distant from tumors, suggesting a localized field effect. Sites with high and low *EGFR* copy numbers were heterogeneously distributed in six of nine primary tumors and in one of eight metastases. EGFR protein overexpression was significantly higher in metastasis sites than in primary tumors. We conclude from our findings that *EGFR* mutations and protein overexpression are early phenomena in the pathogenesis of lung adenocarcinoma and that *EGFR* mutation precedes an increase in gene copy number. In *EGFR*-mutant adenocarcinoma metastases, the higher levels of EGFR overexpression and more homogeneously distributed high gene copy numbers suggest tumor progression. Our findings have important implications for the development of new strategies for targeted chemoprevention and therapy in lung adenocarcinoma using EGFR inhibitors.

Epidermal growth factor receptor (EGFR), a tyrosine kinase (TK) member of the ErbB family, has shown frequent abnormalities in non–small cell lung carcinomas. These abnormalities include protein overexpression, gene amplification, and mutation (1–3). Somatic *EGFR* mutations have been identified in specific subsets of patients with lung adenocarcinoma, including never or light smokers, women, and patients of East Asian descent (4). The mutations cluster in the first four exons (18–21) of the TK domain of the gene, and ~90% of the mutations are composed of either an in-frame deletion in exon 19 or a specific missense mutation in exon 21 (4). An increase in *EGFR* gene copy number, including high polysomy and gene amplification shown by fluorescent *in situ* hybridization (FISH), has been detected in 22% of patients with surgically resected (stages I-IIIA) non–small cell lung carcinomas and correlated with EGFR protein overexpression (2). Higher frequencies (40-50%) of *EGFR* high copy number have been reported in patients with advanced non–small cell lung carcinomas (5–10). Despite this knowledge, limited information is available on the role of EGFR abnormalities in the early pathogenesis and progression of lung adenocarcinomas.

Recently, we showed that mutation of the *EGFR* TK domain is an early event in the pathogenesis of lung adenocarcinoma and is detected in histologically normal bronchial and bronchiolar epithelium (NBE) in 43% of patients with *EGFR*-mutant tumors (11). We found that *EGFR* mutations were more frequent in normal epithelium within the tumor (43%) than in adjacent sites (24%), suggesting a localized field effect (11). However, no comprehensive information is available regarding the role of *EGFR* abnormalities, including gene mutation, increased copy number, and protein overexpression in the early pathogenesis and progression of lung adenocarcinomas.

Both *EGFR* gene mutations and high copy number (gene amplification and high polysomy identified by FISH) have been associated with sensitivity to the small-molecule TK inhibitors gefitinib and erlotinib in patients with lung

---

Authors' Affiliations: Departments of [1]Thoracic/Head and Neck Medical Oncology and [2]Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, and [3]Department of Medicine/Medical Oncology and Pathology, University of Colorado Cancer Center, Aurora, Colorado
Received 02/09/2008; revised 03/28/2008; accepted 04/23/2008.

Grant support: Department of Defense grants W81XWH-04-1-0142 and W81XWH-05-2-0027, the Specialized Program of Research Excellence in Lung Cancer grant P50CA70907, and Cancer Center Support grant CA-16672 from the National Cancer Institute.

Requests for reprints: Ignacio I. Wistuba, Department of Pathology, Unit 85, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030-4009. Phone: 713-563-9184; Fax: 713-792-0309; E-mail: iiwistuba@mdanderson.org.

©2008 American Association for Cancer Research.
doi:10.1158/1940-6207.CAPR-08-0032

adenocarcinoma (5–18). However, some of these results have been rather controversial (9, 10, 19, 20). In these studies of gefitinib and erlotinib, most of the *EGFR* mutation and copy number analyses were done in very small tissue samples or in cytologic specimens obtained from primary tumor and metastasis sites in patients with advanced-stage lung cancer (5–9, 12–16). To date, no studies have been done to identify the characteristics of *EGFR* gene and protein expression abnormalities at different sites with respect to primary lung adenocarcinomas and in corresponding sites of metastasis, information that might resolve some of the controversy.

To identify the sequence of EGFR abnormalities involved in the pathogenesis and progression of lung adenocarcinoma, we did a precise mapping analysis correlating *EGFR* mutation, gene copy number, and protein expression in NBE fields, primary tumors, and corresponding lymph node metastases that were obtained from 50 patients with lung adenocarcinomas, including 24 patients with *EGFR*-mutant primary tumors with nine corresponding lymph node metastasis sites and 26 patients with *EGFR*–wild-type primary tumors.

Materials and Methods

Case selection

To map *EGFR* gene and protein expression abnormalities, we obtained formalin-fixed, paraffin-embedded lung adenocarcinoma tissue specimens from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX). The tumor tissue specimens came from 50 patients with surgically resected lung adenocarcinomas (tumor-node-metastasis stage I-IIIA) with known *EGFR* mutations in exons 18 to 21, as described previously (3, 11). This bank was approved by the M. D. Anderson Cancer Center Institutional Review Board.

Of these 50 patients, 24 patients had lung adenocarcinoma with *EGFR* mutations in exon 18 (n = 1), exon 19 (n = 13), and exon 21 (n = 10), and 26 patients had *EGFR*–wild-type lung adenocarcinoma. The patients' clinicopathologic features are summarized in Table 1. All lung adenocarcinomas were of mixed histologic subtype (WHO classification; ref. 21). None of the patients had received cytotoxic and/or targeted therapy. Clinical staging was based on the revised International System for Staging Lung Cancer (22).

EGFR abnormality mapping

We retrospectively reviewed H&E-stained histology sections of primary tumor, lymph node metastases, and adjacent normal lung tissue specimens to identify tissue foci available for EGFR abnormality analyses. The EGFR abnormalities included *EGFR* mutations in exons 18 and 21, as shown by microdissection and PCR-based sequencing; *EGFR* copy number, as shown by FISH; and total EGFR and phosphorylated EGFR (pEGFR), as shown by immunohistochemical analyses. Representative examples of these molecular changes are illustrated in Fig. 1.

We used serial 5-μm-thick histology sections for the tissue microdissection, FISH, and immunohistochemical analyses. We identified a total of 316 noncontiguous tumor and epithelial foci from among 142 NBE specimens (obtained from 50 patients; 2.84 sites/patient), 144 primary tumors (from 50 patients; 2.88 sites/patient), and 30 lymph node metastases (from 9 patients; 3.3 sites/patient). We examined NBE and primary tumors in both *EGFR*-mutant and *EGFR*–wild-type cases and metastasis sites in *EGFR*-mutant cases only. All epithelial foci consisted of normal or mildly hyperplastic epithelia that harbored small bronchi (65 sites) and bronchioles (77 sites).

The NBE specimens were obtained from three different locations based on their relationship to the tumors: within the tumor (47 sites), ≤5 mm from the tumor margin (adjacent to tumor; 63 sites), and >5 mm from the tumor margin ("distant" lung; 32 sites). We did not detect squamous metaplastic or dysplastic lesions in the bronchial structures or atypical adenomatous hyperplasias in the alveolar tissue. We identified small bronchi on the basis of well-defined smooth muscle and discontinuous cartilage layers. Bronchioles were defined as small conducting airways lacking well-defined smooth muscle wall or cartilage layers. We assessed the location of the small bronchial and bronchiolar respiratory epithelium examined for *EGFR* abnormalities based on the epithelia's location in relation to the tumor tissue in the corresponding histology sections, as previously described (11).

Table 1. Clinicopathologic features of patients with lung adenocarcinomas examined for *EGFR* abnormalities in tumors and adjacent normal epithelium

Features/samples	*EGFR* status		
	Mutant (n = 24)	Wild-type (n = 26)	Total (n = 50)
Mean age (y)	61.3	62.7	62.1
Gender			
Female	19 (79%)	13 (50%)	32
Male	5 (21%)	13 (50%)	18
Ethnicity			
East Asian	13 (54%)	9 (35%)	22
Not East Asian	11 (56%)	17 (65%)	28
Smoking history			
Never	16 (67%)	9 (35%)	25
Former	7 (29%)	10 (38%)	17
Current	1 (4%)	7 (27%)	8
Stage of disease			
I	11 (46%)	15 (58%)	26
II	5 (21%)	4 (15%)	9
IIIA	8 (33%)	7 (27%)	15

Microdissection and DNA extraction

Approximately 1,000 cells were precisely microdissected from 8-μm-thick, H&E-stained, formalin-fixed, paraffin-embedded histology sections for each site using laser capture microdissection (Arcturus Engineering Laser Capture Microdissection System; MDS Analytical Technologies), as previously described (11). To prevent the nonspecific binding of the mutant cells to the microdissection cap film, the microdissected tissue samples were redissected from the film under stereomicroscope visualization using fine needles (25-gauge 5/8-inch needles). We then extracted the DNA using 25 μL of PicoPure DNA Extraction solution containing proteinase K and incubated the DNA at 65°C for 20 h. Subsequently, proteinase K was inactivated by heating samples at 95°C for 10 min.

EGFR mutation analysis

Exons 18 and 21 of *EGFR* were PCR-amplified using DNA extracted from microdissected NBE and tumor cells, as previously described (3, 11). Each PCR was done using HotStarTaq Master Mix (Qiagen) for 40 cycles at 94°C for 30 s, 63°C for 30 s, and 72°C for 30 s, followed by a 7-min extension at 72°C. PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method (Perkin-Elmer Corp.). We confirmed all sequence variants by independent PCR amplifications from at least two independent microdissections and sequenced the variants in both directions.

EGFR FISH analysis

We analyzed the gene copy number per cell using the LSI EGFR SpectrumOrange/CEP 7 SpectrumGreen Probe (Abbott Molecular), as previously described (5). Histology sections were incubated at 56°C overnight and deparaffinized by washing in CitriSolv (Fisher Scientific). After incubation in 2× SSC buffer (pH 7.0) at 75°C for 15 to 25 min, the histology sections were digested with proteinase K (0.25 mg/mL in 2× SSC) at 37°C for 15 to 25 min, rinsed in 2× SSC (pH 7.0) at room temperature for 5 min, and dehydrated using ethanol in a series of increasing concentrations (70%, 85%, 100%). We applied the EGFR SpectrumOrange/CEP 7/SpectrumGreen probe set (Abbott Molecular) onto the selected area, according to the manufacturer's instructions, on the basis of the tumor foci seen on each slide. We then covered the hybridization area with a glass coverslip and sealed the coverslip with rubber cement. The slides were incubated at 80°C for 10 min for codenaturation of chromosomal and probe DNA and then placed in a humidified chamber at 37°C for 20 to 24 h to allow hybridization to occur. Posthybridization washes were done in 1.5 mol/L of urea and 0.1× SSC (pH 7.0-7.5) at 45°C for 30 min and in 2× SSC for 2 min at room temperature. After the samples were dehydrated in a series of increasing ethanol concentrations, 4′,6′-diamidino-2-phenylindole (0.15 mg/mL in Vectashield Mounting Medium; Vector Laboratories) was applied for chromatin counterstaining. FISH analysis was done independently by two authors (M. Varella-Garcia and A. C. Xavier), who were blinded to the patients' clinical characteristics and all other molecular variables. Patients were classified into six FISH strata according to the frequency of cells with the *EGFR* gene copy number and referred to the chromosome 7 centromere, as follows: (*a*) disomy (≥3 copies in <10% of cells); (*b*) low trisomy (3 copies in 10% to 40% of the cells, ≥4 copies in <10% of cells); (*c*) high trisomy (3 copies in ≥40% of cells, ≥4 copies in <10% of cells); (*d*) low polysomy (≥4 copies in 10–40% of cells); (*e*) high polysomy (≥4 copies in ≥40% of cells); and (*f*) gene amplification (ratio of *EGFR* gene to chromosome ≥2, presence of tight *EGFR* gene clusters and 15 copies of *EGFR* per cell in 10% of the analyzed cells). The high polysomy and gene amplification categories were considered to be high *EGFR* copy number, and the other categories were considered to be nonincreased *EGFR* copy number, as previously published (5). Analysis was done in approximately 50 nuclei per tumor and epithelial site, and the section of the area was guided by image captured in the H&E-stained section.

Immunohistochemical staining

Tissue histology sections for immunohistochemical analyses were deparaffinized, hydrated, heated in a steamer for 10 min with 10 mmol/L of sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% bovine serum albumin in TBS with Tween 20 for 30 min at room temperature. For the EGFR analysis, tissue sections were incubated for 2 h with primary antibodies against the EGFR clone 31G7 (1:100 dilution; Zymed) and pEGFR Tyr 1086 (1:100 dilution; Invitrogen). Tissue sections were then incubated for 30 min with the secondary antibody (EnVision+ Dual Link labeled polymer; DAKO), after which diaminobenzidine chromogen was applied for 5 min. The slides were then

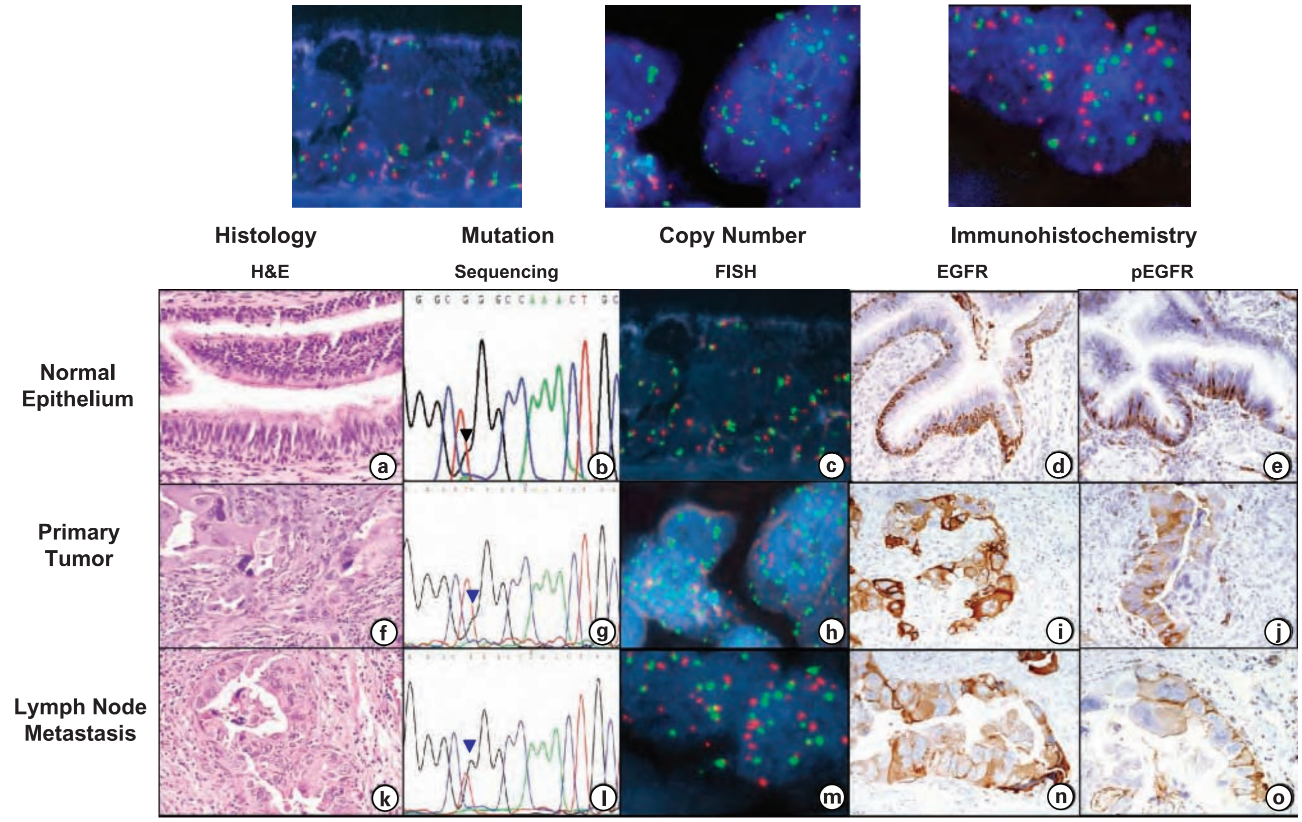


Fig. 1. A representative case of *EGFR*-mutant lung adenocarcinoma: *EGFR* gene and protein expression abnormalities in NBE (*A–E*), primary tumor (*F–J*), and lymph node metastasis (*K–O*) sites. Histologic characteristics (*A*, *F*, and *K*) of tissue sections stained with H&E (magnification, ×100). PCR-based *EGFR* sequencing (*B*, *G* and *L*) of the same *EGFR* mutation in exon 21 (L858R, *black arrowhead*) in NBE (*B*), primary tumor (*G*), and lymph node metastasis sites (*L*). *EGFR* FISH analysis (*C*, *H* and *M*) of low trisomy (low copy number) in the NBE sample (*C*), high polysomy in the primary tumor site (*H*), and gene amplification (*M*) in the metastasis site. Immunohistochemical analysis (*D*, *I*, *N*, *E*, *J*, and *O*) of high EGFR and pEGFR expression in the membrane and cytoplasm in all three types of samples.

Table 2. Frequency of *EGFR* gene mutation and protein overexpression in histologically normal bronchial and bronchiolar epithelium obtained from *EGFR*-mutant and wild-type lung adenocarcinomas

EGFR abnormality in NBE	Cases			Sites		
	Mutant	Wild-type	Total	Mutant	Wild-type	Total
Mutation by sequencing						
Number	24	26	50	85	57	142
Mutant	12 (50%)[*]	3 (12%)[*]	15 (30%)	22 (26%)	8 (14%)	30 (21%)
Protein overexpression by immunohistochemistry[†]						
Number	23	26	49	78	56	134
EGFR	19 (83%)	15 (58%)	34 (69%)	52 (67%)	35 (63%)	87 (65%)
pEGFR	10 (44%)	6 (23%)	16 (33%)	24 (31%)	12 (21%)	36 (27%)

[*] $P = 0.003$.

[†] Positive immunohistochemical overexpression score >200 (range, 0-400).

counterstained with hematoxylin and topped with a coverslip. For EGFR and pEGFR expression, antibody specificity was confirmed using blocking peptide and phosphatase incubation experiments. For the control experiments, we used formalin-fixed and paraffin-embedded pellets from lung cancer cell lines with confirmed EGFR and pEGFR overexpression. Thyroid transcription factor-1 (TITF-1) antibody (1:100 dilution, Cell Marque) was used for the identification of TITF-1–positive cells. All four antibodies were incubated for 1.5 h at room temperature. Immunohistochemistry results were scored jointly by two authors (X. Tang and I.I. Wistuba), who were blinded to clinical and other molecular variables. Immunostaining of the cell membrane and cytoplasm for EGFR and pEGFR was evaluated by light microscopy (magnification, ×20). A semiquantitative approach was used to generate a score for each tissue site, as previously described (2, 23, 24). Membrane and cytoplasm stains were recorded separately. We defined the intensity score as follows: 0, no appreciable staining in the NBE or malignant cells; 1, barely detectable staining in NBE or malignant cells compared with the stromal elements; 2, readily appreciable staining; 3, dark brown staining of cells; and 4, very strong staining of cells. The score was also based on the fraction of cells showing a given staining intensity (0-100%). We calculated the immunohistochemical scores by multiplying the intensity and extension, and the scores ranged from 0 to 400. For the statistical analyses, scores of 0 to 200 signified negative/low expression, and scores >200 indicated positive/overexpression, as previously reported (2, 23, 24). For the evaluation of nuclear TITF-1 immunohistochemical expression, 200 epithelial cells were quantified by light microscopy (magnification, ×20), and a score (range, 0-100) expressing the percentage of positive cells was obtained.

Statistical analysis

All relationships between categorical variables were assessed using χ^2 and Fisher's exact tests. $P < 0.05$ values were considered statistically significant.

Table 3. *EGFR* mutation and protein overexpression in histologically normal epithelium by location

EGFR abnormality in NBE	Location in relation to the tumor			Structure	
	Inside	Adjacent	Distant	Bronchiole	Small bronchus
Mutation					
Mutant tumor	11/31 (36%)[*]	10/35 (29%)	1/17 (6%)[*]	10/43 (23%)	12/42 (29%)
Wild-type tumor	2/15 (13%)	3/28 (11%)	1/15 (7%)	4/34 (12%)	2/23 (9%)
All tumors	13/46 (28%)	13/63 (21%)	2/32 (6%)	14/77 (19%)	14/65 (22%)
EGFR overexpression[†]					
Mutant tumor	24/29 (83%)[‡]	20/33 (61%)[‡]	8/16 (50%)[‡]	18/38 (47%)[§]	34/40 (85%)[§]
Wild-type tumor	10/15 (67%)	17/28 (61%)	8/13 (62%)	14/33 (42%)[§]	21/23 (91%)[§]
All tumors	34/44 (77%)	37/61 (61%)	16/29 (55%)	32/71 (45%)[§]	55/63 (87%)[§]
pEGFR overexpression[†]					
Mutant tumor	13/29 (45%)[‖]	5/33 (15%)[‖]	6/16 (38%)[‖]	10/38 (26%)	14/40 (35%)
Wild-type tumor	5/15 (33%)	5/28 (18%)	2/13 (15%)	2/33 (6%)[§]	10/23 (44%)[§]
All tumors	18/44 (41%)	10/61 (16%)	8/29 (28%)	12/71 (17%)[¶]	24/63 (38%)[¶]

[*] Comparison of NBE from inside tumor vs. NBE distant ($P = 0.02$).

[†] Positive immunohistochemical overexpression score >200 (range 0-400).

[‡] Comparison of NBE from inside tumor vs. NBE adjacent + distant ($P = 0.02$)

[§] Comparison of NBE from bronchiole vs. small bronchus ($P < 0.001$).

[‖] Comparison of NBE from inside tumor vs. NBE adjacent + distant ($P = 0.038$).

[¶] Comparison of NBE from bronchiole vs. small bronchus ($P = 0.006$).

Results

EGFR abnormalities in the early pathogenesis of lung adenocarcinomas

Patterns of EGFR mutation in NBE. We previously reported our finding of mutations in exons 19 and 21 of *EGFR* in at least one site of microdissected NBE obtained from lung cancer specimens from 9 of 21 (43%) patients with *EGFR*-mutant adenocarcinomas, with no such mutations found in any of 26 respiratory epithelium foci from 16 patients with wild-type tumors (11). In the present study, using the same methodology, we analyzed for *EGFR* mutation in NBE obtained from an additional 3 patients with an *EGFR*-mutant and 10 patients with *EGFR*–wild-type lung adenocarcinomas. Combining both data sets, the overall rate of mutation in NBE from *EGFR*-mutant tumors was 50%. In the wild-type tumor cases, we detected *EGFR* exon 19 deletions (15 bp, 746-750) in six sites of small bronchial (*n* = 4) and bronchiolar (*n* = 2) NBE obtained from three wild-type tumors (Table 2). Thus, an *EGFR* mutation was found in NBE in 3 of 26 (12%) wild-type adenocarcinomas and in 8 of 57 (14%) of the microdissected epithelial sites (Table 2).

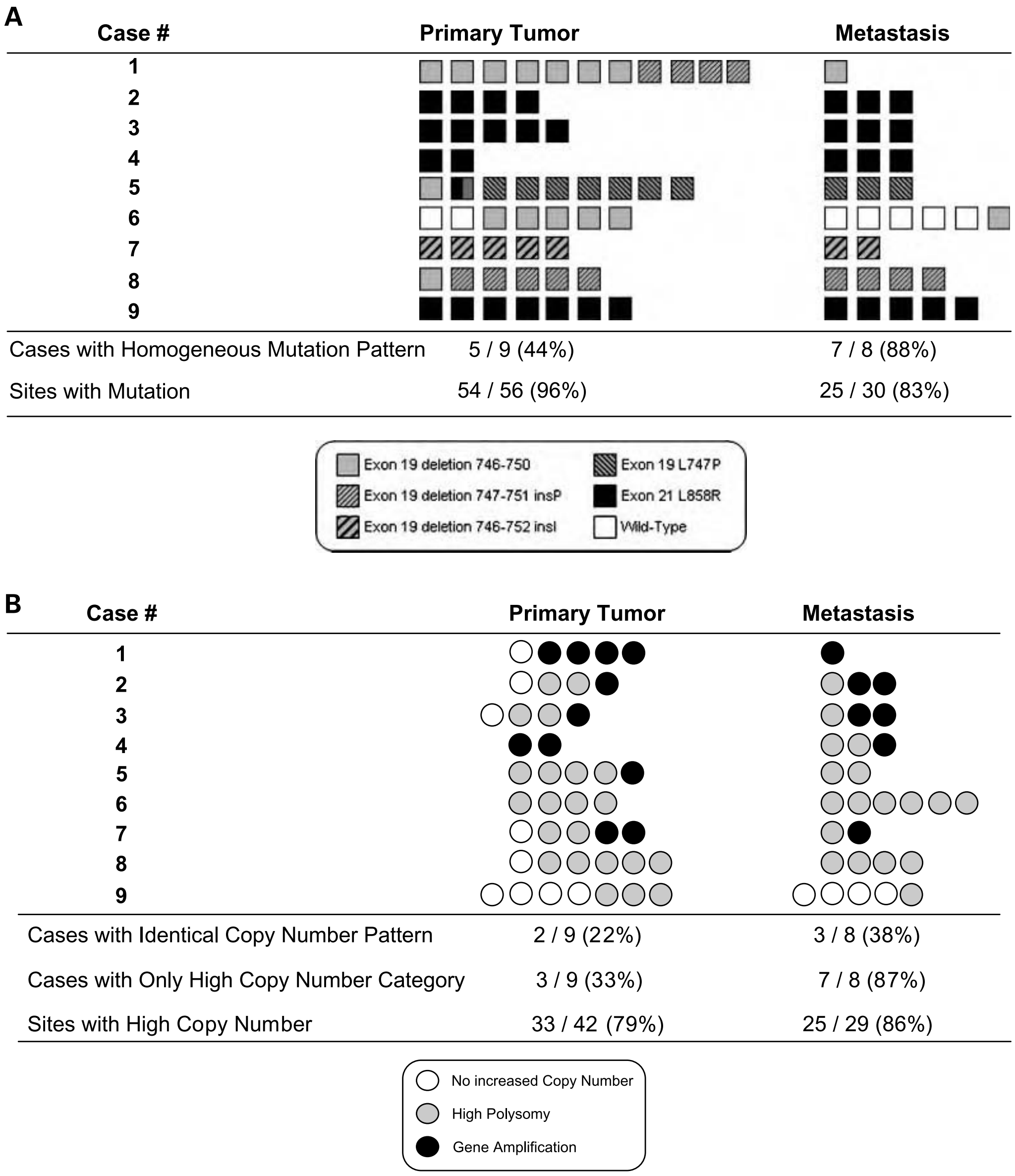


Fig. 2. *A*, *EGFR* mutation pattern in 56 primary tumor and 30 lymph node metastasis sites obtained from nine patients with *EGFR*-mutant lung adenocarcinomas. A homogeneous mutation pattern was detected in five primary tumors (*cases 2*, *3*, *4*, *7*, and *9*) and all but one (*case 6*) metastasis case. *Case 6*, mixed wild-type and mutant sites in both primary tumor sites and corresponding metastases. *B*, *EGFR* copy number pattern shown by FISH in 42 primary tumor and 29 lymph node metastasis sites obtained from nine patients with *EGFR*-mutant lung adenocarcinomas. Different FISH copy number categories (low vs. high) were found in six of nine primary tumors and in one of eight corresponding metastases. Positive *EGFR* FISH expression included high polysomy and gene amplification, and negative *EGFR* FISH expression included disomy and trisomy.

Table 4. Summary of *EGFR* abnormalities by sites in nine primary lung adenocarcinomas and corresponding lymph node metastases

***EGFR* abnormality/ number of sites**	**Primary tumor**	**Metastases**
Mutation		
Number of sites examined	56	30
Mutation positive	54 (96%)*	25 (83%)*
Copy no.		
Number of sites examined	42	29
Low copy no.	9 (21%)	4 (14%)
High copy no.	33 (79%)	25 (86%)
High polysomy	22 (52%)	18 (62%)
Gene amplification	11 (26%)	7 (24%)
Protein overexpression†		
Number of sites examined	65	31
EGFR	42 (65%)‡	30 (97%)‡
pEGFR	9 (14%)§	21 (68%)§

*The same case harbored two primary tumor and five metastasis sites with *EGFR*-wild-type sequence.

†Positive immunohistochemical expression score >200 (range 0-400).

‡Primary tumor vs. metastasis ($P = 0.02$).

§Primary tumor vs. metastasis ($P = 0.00001$).

The combined data showed that NBE with mutant *EGFR* was detected in the small bronchi (13 of 64, 20%) and bronchioles (17 of 78, 22%) of both mutant and wild-type tumor cases. Overall, however, the mutation frequency was higher in NBE samples microdissected from within the tumor (13 of 47, 28%) than in samples obtained from adjacent tissue and tissue distant from the tumors (17 of 95, 18%; Table 3).

In our previously reported comparison of NBE and corresponding tumors (16 specimens), we always observed identical *EGFR* mutations in both sites examined (11). In this study, we have expanded the number of NBE sites ($n = 85$) examined for the mutation in patients with *EGFR*-mutant adenocarcinomas and detected five sites (6%) from three cases in which NBE showed mutations different from the ones detected in the corresponding tumor specimens (data not shown). Importantly, in all cases with a mutation in NBE, an identical mutation was detected in at least one site of the corresponding tumor specimen. Thus, in this expansion of our previous study (11), a relatively more heterogeneous *EGFR* mutation pattern of the respiratory field was detected in NBE microdissected from mutant lung adenocarcinomas, but most NBE and corresponding tumors shared the same mutation.

***EGFR* copy number and correlation with gene mutation in NBE.** To determine the morphologic stage at which *EGFR* copy abnormalities arise in *EGFR*-mutant adenocarcinomas, we did a precise mapping analysis and examined *EGFR* copy number in 21 NBE sites obtained from nine mutant adenocarcinomas using FISH. All nine tumor specimens showed at least one site with a high copy number. These epithelial sites were also examined in the *EGFR* mutation analysis. Most NBE (14 of 21, 67%) showed no *EGFR* FISH abnormalities (disomy), including four *EGFR*-mutant sites with exon 19 (15 bp) deletions and exon 21 (L858R) point mutations. Trisomy was detected in seven (33%) NBE sites obtained from six (67%) cases. We did not identify any NBE with *EGFR* amplifications or a high level of polysomy, which have been defined as high gene copy number. In contrast, the nine tumors mapped showed significantly higher frequency of *EGFR* amplification (11 of 42 sites, 26%; $P < 0.018$) or a high level of polysomy (22 of 42, 52%; $P < 0.001$) compared with NBE. Our findings indicate that high *EGFR* copy number does not occur in peripheral NBE in *EGFR*-mutant lung adenocarcinomas and that gene mutations precede copy number abnormalities in the sequential pathogenesis of these tumors.

***EGFR* immunohistochemical expression and correlation with gene mutation in NBE.** We evaluated the level of EGFR and pEGFR protein expression in 134 NBEs obtained from *EGFR*-mutant and wild-type lung adenocarcinomas. Overall, a

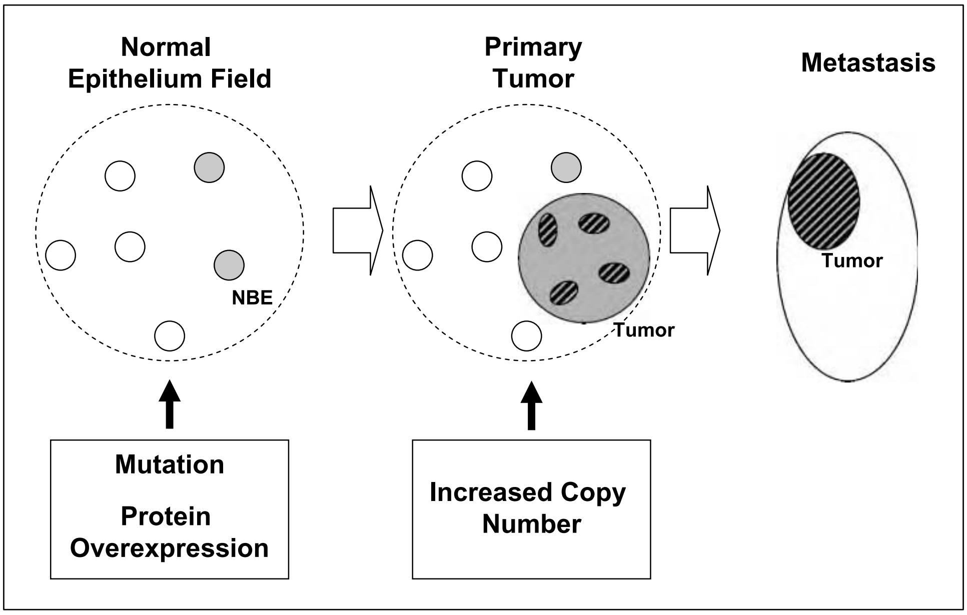


Fig. 3. Proposed sequence of *EGFR* abnormalities occurring in the early pathogenesis and progression of *EGFR*-mutant lung adenocarcinomas. NBE field, primary tumor, and metastasis sites. *Small circles*, NBE, which acquires *EGFR* mutations and EGFR protein (total and phosphorylated) overexpression (*gray circles*). In the primary tumor stage, the *EGFR* copy number increases (high polysomy and gene amplification) in small tumor foci (*striped ovals*). In the metastasis site, tumor cells show both *EGFR* mutation and high copy number throughout most of the lesion.

high level of EGFR (69%) and a moderate level of pEGFR (33%) expression were detected in NBE from patients with tumors (Table 2). However, EGFR and pEGFR were expressed to a greater degree in NBE sites obtained from patients with *EGFR*-mutant tumors than in patients with wild-type tumors (Table 2), although these differences were not statistically significant. The frequency of EGFR, but not of pEGFR, overexpression was higher in *EGFR*–wild-type NBE sites (85 of 111, 77%) than in mutant sites (14 of 24, 58%; $P = 0.039$). Of interest, NBE located inside tumors showed the highest frequency of EGFR and pEGFR overexpression compared with NBE located adjacent to and distant from tumors, especially in *EGFR*-mutant tumors (Table 3). Small bronchi also showed a higher frequency of overexpression of both markers compared with bronchioles (Table 3). Thus, the overexpression of EGFR and pEGFR is a common event in NBE from patients with lung adenocarcinomas, especially in *EGFR*-mutant tumors, and shows a localized field phenomenon effect similar to gene mutation.

TITF-1 immunohistochemical expression and* EGFR *mutation in NBE. Recently, on the basis of immunohistochemical findings of higher levels of nuclear TITF-1 expression, a crucial transcription factor of the lung, in *EGFR*-mutant lung adenocarcinomas compared with in wild-type tumors, it has been suggested that *EGFR*-mutant lung adenocarcinoma originates from the terminal respiratory unit (25), which is composed of alveolar cells and nonciliated bronchiolar epithelium. Its characteristics are highlighted by the expression of TITF-1 (25). We therefore investigated the correlation between *EGFR* mutation and TITF-1 nuclear expression in tumor and normal epithelium sites. *EGFR*-mutant lung adenocarcinomas (18 of 20 cases, 90%) showed higher expression of TITF-1 than did wild-type tumors (10 of 26 cases, 38%; $P < 0.001$). However, in immunohistochemical studies, we did not see a significant difference in the frequency of TITF-1 expression between *EGFR*-mutant (11 of 25 sites, 44%) and *EGFR*–wild-type (34 of 105 sites, 33%; $P = 0.273$) respiratory epithelia. Our findings therefore indicate that NBE cells expressing TITF-1 are not the exclusive precursors of *EGFR*-mutant adenocarcinomas. From these results, it is clear that these tumors do not originate exclusively from terminal respiratory unit structures.

EGFR abnormalities in the progression of lung adenocarcinomas

EGFR mutation pattern in primary tumors and corresponding metastasis. To identify the characteristics of EGFR abnormalities in the progression of mutant lung adenocarcinomas, we examined *EGFR* gene mutation, gene copy number, and protein expression in primary tumors and corresponding metastases by performing a detailed mapping analysis of tumor specimens. For this study, we selected nine lung adenocarcinomas with known *EGFR* mutations in exon 19 ($n = 5$) and exon 21 ($n = 4$), and with lymph node metastases for which there was sufficient tissue to perform our mapping analysis.

For the mutation analysis of *EGFR* exons 19 and 21, we did precise tissue microdissection from noncontiguous primary tumor foci ($n = 56$ sites, 6.2 sites/tumor; range 2-11 sites) containing at least 1,000 cells. Surprisingly, four of the nine primary tumorsexamined showed mixed *EGFR* gene patterns (Fig. 2A): three showed two or more types of mutations, and one showed five sites with exon 19 (15 bp, 746-750) deletion and two sites with the wild-type *EGFR* gene. *EGFR* mutation analysis of 30 corresponding lymph node metastasis sites from the nine *EGFR*-mutant cases (3.3 sites/case; range 1-6 sites) detected only one type of *EGFR* mutation in all tumor sites in each case, and the mutation was always present in at least one site of the corresponding primary tumor. Similar to the corresponding primary tumor, one metastasis case showed *EGFR*–wild-type (five sites) and *EGFR*-mutant [one site, exon 19 (15 bp, 746-750) deletion] tumor sites (Fig. 2A). All these findings were confirmed by sequencing analyses of independently microdissected samples. In summary, our findings showed a relatively high level of heterogeneity for the *EGFR* mutation, and several tumor cell clones had mutation patterns in the primary tumor specimens that differed from the mutation patterns in the lymph node metastasis sites.

EGFR copy number abnormalities in primary tumors and corresponding metastasis. We used FISH to investigate the *EGFR* gene copy number abnormalities in 42 primary tumor sites (2.1 sites/case; range 2-7 sites) and 29 metastasis sites (3.2sites/case; range 1-6 sites), which were also examined for the mutation analysis. Overall, all primary tumors and corresponding metastases showed at least one site of high gene copy number (high polysomy or gene amplification; Fig. 2B). However, six (67%) primary tumor cases and one (11%) metastasis case showed at least one site without high copy number (disomy in one primary tumor site, high trisomy in one metastasis site, and low polysomy in seven primary and three metastasis sites; Fig. 2B). Thus, *EGFR* copy number heterogeneity was higher in primary tumor sites than in corresponding metastasis sites.

EGFR immunohistochemical expression in primary tumors and corresponding metastasis sites. In the nine *EGFR*-mutant lung adenocarcinoma cases mapped for *EGFR* abnormalities, we examined both primary tumors and the corresponding lymph node metastases for EGFR and pEGFR immunohistochemical expression. For both tumor locations combined, 96 distinct tumor sites were examined ($n = 65$ primary tumor sites, 7.2 sites/case; and $n = 31$ metastasis sites, 3.4 sites/case). Significantly higher levels of EGFR and pEGFR expression were detected in metastasis sites compared with primary tumor sites (Table 4). No correlation between EGFR and pEGFR expression and *EGFR* copy number status by FISH was detected.

Discussion

Using a detailed molecular pathology mapping strategy, we determined the sequence of EGFR abnormalities in the early pathogenesis of *EGFR*-mutant lung adenocarcinomas and identified the pattern of EGFR changes in the progression of *EGFR*-mutant lung adenocarcinomas from primary tumors to lymph node metastasis. First, we showed that *EGFR* mutations precede gene copy number abnormalities in the pathogenesis of these tumors and that EGFR and pEGFR immunohistochemical protein expressions are frequent events in histologically normal peripheral bronchial and bronchiolar epithelium adjacent to lung adenocarcinomas. Second, our data indicated that although primary lung adenocarcinomas show some degree of *EGFR* gene copy number heterogeneity, this phenomenon is rare in metastases. Although these findings can be considered tumor progression phenomena, they also have important clinical implications from the standpoint of making decisions regarding the use of EGFR TK inhibitor therapy on the basis of the finding of *EGFR* gene abnormalities.

Despite the evidence showing that atypical adenomatous hyperplasia is a precursor of peripheral lung adenocarcinomas (26), there is consensus that the pathogenesis of most adenocarcinomas is unknown. Our previously reported findings of an *EGFR* mutation in NBE in 9 of 21 (43%) patients with *EGFR*-mutant adenocarcinomas indicated that the *EGFR* gene mutation is an early event in the pathogenesis of lung adenocarcinoma (11). In this study, we have investigated normal epithelium from additional patients with *EGFR*-mutant or wild-type lung adenocarcinomas and specifically have two new findings in this study; (*a*) we detected an *EGFR* mutation (exon 19, 15 bp deletion, 746-751) in six sites of small bronchial and bronchiolar epithelium obtained from three patients with wild-type adenocarcinoma, and (*b*) whereas an identical mutation was detected in the majority of specimens of mutant normal epithelium compared with the corresponding invasive tumor (75% of cases and 94% of sites), we found few normal epithelium sites (6%) in three of 12 cases (25%) of *EGFR*-mutant tumors, demonstrating the existence of a different mutation pattern between normal epithelium and the corresponding invasive tumor. All these data reinforce the concept of a field effect phenomenon in *EGFR* mutations in lung adenocarcinoma pathogenesis that affects histologically normal bronchial and bronchiolar respiratory epithelia.

We have previously shown that molecular abnormalities occur in a stepwise fashion in the sequential pathogenesis of squamous cell carcinoma of the lung, with molecular changes commencing in histologically normal bronchial epithelium in smokers and in patients with lung cancer (27, 28). Our findings suggest that *EGFR* abnormalities also occur sequentially in the early pathogenesis of lung adenocarcinoma, with a mutation commencing in histologically normal epithelium and a high *EGFR* copy number appearing at the invasive tumor stage. A recent report (29) of selective gene amplification of the shorter allele of the *EGFR* intron 1 polymorphism CA simple sequence repeat 1, which is the allele more frequently mutated in tumors harboring an *EGFR* mutation, also suggests that mutations occur earlier than copy number abnormalities in the pathogenesis of lung adenocarcinoma. Our findings of frequent EGFR (69%) and pEGFR (33%) protein overexpression in normal distal bronchial and bronchiolar epithelium from patients with either *EGFR*-mutant or *EGFR*–wild-type lung adenocarcinomas indicate a field phenomenon in the peripheral airway. A relatively high frequency of EGFR protein expression has also been reported in centrally located, histologically normal (42%) and hyperplastic (54%) bronchial epithelium from smokers (23). In addition, our data indicate that the mechanisms of protein overexpression seem to be unassociated with high gene copy number and mutation in NBE. Other mechanisms can explain EGFR overexpression in normal epithelial cells, including ligand-dependent up-regulation and activation, as well as inhibition of endocytosis-related protein down-regulation in the cell membrane (30).

Based on findings of higher levels of immunohistochemical expression of nuclear TITF-1, a crucial transcription factor of the lung, in *EGFR*-mutant lung adenocarcinomas compared with wild-type tumors, it has been suggested that *EGFR*-mutant lung adenocarcinoma originates from the terminal respiratory unit (25). We found *EGFR* mutations in microdissected histologically normal epithelial cells from small bronchi and bronchioles, which supports the concept of adenocarcinomas arising from the peripheral lung airway. Our findings indicate that NBE cells expressing TITF-1 are not the exclusive precursors of *EGFR*-mutant adenocarcinomas. From this finding, it is clear that these tumors do not originate exclusively from terminal respiratory unit structures. In addition, we cannot exclude the possibility that common stem or progenitor cells for both bronchial and bronchiolar epithelium bear *EGFR* mutations.

It has been suggested that activating TK *EGFR* mutations are a potent oncogenic event by which mutant tumor cells become physiologically dependent on the continued activity of the phosphorylated protein for the maintenance of their malignant phenotype (31). Our detailed mapping analysis of the *EGFR* gene mutation and copy number of multiple precisely microdissected sites in nine mutant primary tumors and corresponding lymph node metastases showed an identical or monoclonal pattern of mutation in most (n = 5) primary tumors and all metastases. These findings corroborate the monoclonal concept of tumor development and the monoclonal evolution of metastases (32, 33). However, two primary tumors lacking identical or monoclonal *EGFR*-mutant patterns harbored different sizes of exon 19 deletions (12 versus 15 bp and 15 versus 18 bp deletions). This finding could be explained by a tumor progression phenomenon in which the deletion size changed during the evolution of the malignancy. However, two very interesting primary tumors in our study exhibited findings that challenged the concept of the monoclonal evolution of tumors. One case showed a single site with an exon 19 (15 bp) deletion, whereas the remaining eight sites lacked the deletion but showed a point mutation (TTA747CCA) in the same exon. Of interest, the three metastasis sites examined harbored the most frequent mutation detected in the primary lung tumor. The other case showed areas of wild-type and mutant *EGFR* in both primary tumors and metastases, a phenomenon that is difficult to explain and suggests that molecular events other than an *EGFR* mutation may be responsible for tumor development in lung adenocarcinomas. These findings were confirmed by the sequencing of independently microdissected samples. In the latter case, the finding of a high *EGFR* copy number (high polysomy) in wild-type tumor sites raises the possibility of an alternative explanation—that the wild-type allele is preferentially amplified in some tumor cells. As a result, the mutant allele is underrepresented and is not detectable by our current sequencing methodology.

Retrospective studies have provided data suggesting that a high *EGFR* gene copy number shown by FISH is associated with treatment response, time to progression, and survival in patients with advanced non–small cell lung carcinoma treated with EGFR TK inhibitors (5-7, 10, 17). In these studies, high *EGFR* copy number as shown by FISH was defined as true gene amplification or high polysomy with equal to or more than four *EGFR* copies in ≥40% of cells (5, 34). Our mapping analysis of primary tumors and corresponding lymph node metastases in which we used the same *EGFR* FISH criteria showed that a frequent high copy number in mutant tumors was the most frequent pattern detected. Despite the fact that most primary tumor sites and nearly all metastasis sites showed high copy numbers, high polysomy and gene amplification were heterogeneously distributed in both tumor locations. More importantly, five of nine (56%)

primary tumors and one metastasis (13%) showed one or more sites without an increased copy number (FISH negative). Similarly, EGFR and pEGFR immunohistochemical expression was less heterogeneous in primary tumors and more frequent in metastases. Taken together, these data suggest that *EGFR* copy number analyzed by FISH and protein expression analyzed by immunohistochemistry in small core biopsy or fine-needle aspiration specimens obtained from primary tumors, and more rarely from metastases, could miss these molecular changes, especially if only a small number of malignant cells are available for examination. In addition, if the suggested presence of *EGFR* high copy number correlates with sensitivity to EGFR TK inhibitors (5–7, 17), it is likely that metastases will show a better response to therapy than will primary tumors. This is an important consideration, in light of the fact that the site of origin (primary versus metastasis) of the tumor specimen was not reported and factored into the biomarker analyses in any of the clinical trials testing the efficacy of EGFR TK inhibitors in patients with advanced non–small cell lung carcinoma in whom *EGFR* copy number determined by FISH was examined as a predictor of response and prognosis (5–7). Our results show that a better understanding of the pattern of molecular abnormalities and their corresponding biomarker expression, including primary tumors and the frequent metastases seen for this tumor type, is important in lung cancer.

In summary, our data suggest that gene mutations and protein overexpression are the earliest phenomena in *EGFR*-mutant lung adenocarcinoma, occurring at the NBE stage, and that this is followed by the development of a focal increase in copy number at the tumor stage (Fig. 3). At the metastasis sites, however, all three abnormalities were more frequent than they were in the primary tumors and were homogeneously distributed throughout the malignant cells.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR. The biology of epidermal growth factor receptor in lung cancer. Clin Cancer Res 2004;10:4227s–32s.
2. Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol 2003;21:3798–807.
3. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97:339–46.
4. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006;118:257–62.
5. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. J Natl Cancer Inst 2005;97:643–55.
6. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 2005;353:133–44.
7. Hirsch FR, Varella-Garcia M, McCoy J, et al. Increased epidermal growth factor receptor gene copy number detected by fluorescence *in situ* hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. J Clin Oncol 2005;23:6838–45.
8. Jackman DM, Holmes AJ, Lindeman N, et al. Response and resistance in a non-small-cell lung cancer patient with an epidermal growth factor receptor mutation and leptomeningeal metastases treated with high-dose gefitinib. J Clin Oncol 2006;24:4517–20.
9. Massarelli E, Varella-Garcia M, Tang X, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Clin Cancer Res 2007;13:2890–6.
10. Bunn PA, Jr., Dziadziuszko R, Varella-Garcia M, et al. Biological markers for non-small cell lung cancer patient selection for epidermal growth factor receptor tyrosine kinase inhibitor therapy. Clin Cancer Res 2006;12:3652–6.
11. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65:7568–72.
12. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.
13. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
14. Mitsudomi T, Kosaka T, Endoh H, et al. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. J Clin Oncol 2005;23:2513–20.
15. Han SW, Kim TY, Hwang PG, et al. Predictive and prognostic impact of epidermal growth factor receptor mutation in non-small-cell lung cancer patients treated with gefitinib. J Clin Oncol 2005;23:2493–501.
16. Taron M, Ichinose Y, Rosell R, et al. Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. Clin Cancer Res 2005;11:5878–85.
17. Han SW, Kim TY, Jeon YK, et al. Optimization of patient selection for gefitinib in non-small cell lung cancer by combined analysis of epidermal growth factor receptor mutation, K-ras mutation, and Akt phosphorylation. Clin Cancer Res 2006;12:2538–44.
18. Pugh TJ, Bebb G, Barclay L, et al. Correlations of EGFR mutations and increases in EGFR and HER2 copy number to gefitinib response in a retrospective analysis of lung cancer patients. BMC Cancer 2007;7:128.
19. Sequist LV, Haber DA, Lynch TJ. Epidermal growth factor receptor mutations in non-small cell lung cancer: predicting clinical response to kinase inhibitors. Clin Cancer Res 2005;11:5668–70.
20. Johnson BE, Janne PA. Selecting patients for epidermal growth factor receptor inhibitor treatment: a FISH story or a tale of mutations? J Clin Oncol 2005;23:6813–6.
21. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: W.D. Travis E, Brambilla HK, Muller-Hermelink, C.C. Harris, editors. Pathology and genetics: Tumours of the lung, pleura, thymus and heart. Lyon: IARC; 2004. p. 9-124.
22. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111:1710–7.
23. Merrick DT, Kittelson J, Winterhalder R, et al. Analysis of c-ErbB1/epidermal growth factor receptor and c-ErbB2/HER-2 expression in bronchial dysplasia: evaluation of potential targets for chemoprevention of lung cancer. Clin Cancer Res 2006;12:2281–8.
24. Tsao AS, Tang XM, Sabloff B, et al. Clinicopathologic characteristics of the EGFR gene mutation in non-small cell lung cancer. J Thorac Oncol 2006;1:231–9.
25. Yatabe Y, Kosaka T, Takahashi T, Mitsudomi T. EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. Am J Surg Pathol 2005;29:633–9.
26. Westra WH. Early glandular neoplasia of the lung. Respir Med 2000;1:163–9.
27. Wistuba II, Behrens C, Milchgrub S, et al. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. Oncogene 1999;18:643–50.
28. Wistuba II, Behrens C, Virmani AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.
29. Nomura M, Shigematsu H, Li L, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007;4:e125.
30. Grandal MV, Madshus IH, Epidermal growth factor receptor and cancer: control of oncogenic signalling by endocytosis. J Cell Mol Med 2008 [Epub ahead of print].
31. Gazdar AF, Shigematsu H, Herz J, Minna JD. Mutations and addiction to EGFR: the Achilles 'heal' of lung cancers? Trends Mol Med 2004;10:481–6.
32. Fearon ER, Hamilton SR, Vogelstein B. Clonal analysis of human colorectal tumors. Science 1987;238:193–7.
33. Garcia SB, Novelli M, Wright NA. The clonal origin and clonal evolution of epithelial tumours. Int J Exp Pathol 2000;81:89–116.
34. Varella-Garcia M. Stratification of non-small cell lung cancer patients for therapy with epidermal growth factor receptor inhibitors: the EGFR fluorescence *in situ* hybridization assay. Diagn Pathol 2006;1:19.

Chromosomal Deletions and Progression of Premalignant Lesions: Less Is More

Ignacio I. Wistuba[1] and Matthew Meyerson[2,3]

Epithelial malignancies arise after a series of progressive pathologic changes including hyperplasia, different grades of dysplasia, and carcinoma *in situ* (1). These premalignant changes can accompany cancer or may occur in the epithelium of individuals at high risk. Sequential premalignant lesions have been defined for many epithelial tumors, including oral squamous cell carcinoma (1) and Barrett's esophageal adenocarcinoma (3). Two of the most exciting areas of current cancer research are (a) novel high-throughput technology for molecular studies of carcinogenesis and (b) the use of this technology in discovering and characterizing genetic abnormalities that underlie the progression of epithelial premalignancy.

Many studies of the last 2 decades have helped to characterize molecular changes of epithelial premalignancy. Encouraged by the development of methodologies (such as laser microdissection) for isolating cells from small histologic lesions, and of techniques to do genomic studies on minute amounts of DNA, several groups have made substantial progress in unveiling the molecular and genetic abnormalities of epithelial premalignant lesions (2, 3–7). These changes involve inactivation of known and putative tumor suppressor genes and several dominant oncogenes. Tumor suppressor genes are believed to be inactivated via a two-step process involving both alleles. Knudson (8) proposed that the first “hit” frequently is a point mutation, whereas the second allele is subsequently inactivated via a chromosomal deletion, translocation, or other event such as methylation of gene promoter regions (9). Dominant oncogenes are frequently activated by mutation, increased copy number, and translocations (9).

The general working model of sequential molecular abnormalities in the pathogenesis of epithelial tumors indicates that genetic changes (*a*) commence in histologically normal epithelium and progress with the increasing severity of epithelial changes, (*b*) follow a sequence from early to late changes, and (*c*) are extensive and multifocal throughout the epithelium, indicating a field effect, also known as “field cancerization” (3). Therefore, in various organ sites, multiple clonal and subclonal patches of molecularly abnormal epithelial cells (with or without cytologic and morphologic abnormalities) can be detected throughout the affected epithelium. New high-throughput genomic and proteomic profiling techniques are now starting to be applied to premalignant or normal-appearing tissue because these techniques are suitable for studying the small amounts of tissue usually available in these precancer settings.

Oral squamous cell carcinoma and Barrett's esophageal adenocarcinoma are on the frontier of innovative discoveries involving molecular events in the progression of premalignant lesions (2, 3). These two diseases are good models of cancer genetic progression because the premalignant epithelium can be safely visualized and biopsied so that genomic changes can be compared in different stages of neoplastic evolution and then studied longitudinally by biopsy surveillance. Several studies have shown that the general working model of progression described above applies to both tumor types.

Oral leukoplakia is the most common head and neck premalignant lesion and has a malignant transformation rate as high as 24% (10). Deletion of one of the two alleles within chromosome arm 3p and chromosome segment 9p21 is the most frequent event in oral lesions with only mild dysplasia and even in some normal-appearing epithelial cells (11). Oral leukoplakias with deletions in the 3p14 (*FHIT*) and 9p21 (*CDKN2A*) regions are reported to carry a higher risk for transformation (2). Barrett's esophagus (BE) is the only known premalignancy of esophageal adenocarcinoma; only a small fraction (0.5-1% annually) of BE patients subsequently develop adenocarcinoma (12). Although less accessible than are oral lesions, BE is a unique model for the study of human neoplastic progression *in vivo*. The standard care of BE requires biopsies according to defined protocols at multiple time points from the same patient, allowing the generation of spatial maps and longitudinal evaluation of genetic alterations that arise during clonal evolution. Several studies of BE indicate that inactivation of *CDKN2A* by loss of heterozygosity (LOH), methylation, and/or mutation is selected as an early event that predisposes to large clonal expansions of the BE tissue (12, 13). Subsequent inactivation of *TP53* by mutation and LOH predisposes to progression to aneuploidy and invasive adenocarcinoma development (13). All these studies have been done using low-throughput methodologies with analysis of abnormalities on few chromosomal foci or specific genes.

In the present issue of the journal, Tsui et al. (14) and Li et al. (15) report their high-throughput analyses of genetic abnormalities in premalignant lesions that provide insights on chromosomal changes in the early pathogenesis of oral squamous cell carcinoma (14) and BE (15). These studies are among the first to use high-throughput DNA copy number analysis by microarrays in the study of the sequential progression of premalignant lesions of any organ site. Comparative genomic

---

Authors' Affiliations: [1]Departments of Pathology and Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas; [2]Department of Medical Oncology and Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, Massachusetts; and [3]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts

Received 09/17/2008; accepted 09/29/2008.

Grant support: U.S. Department of Defense, VITAL, grant W81XWH-04-1-0142 (I.I. Wistuba).

Requests for reprints: Ignacio I. Wistuba, Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Unit 85, 1515 Holcombe Boulevard, Houston, TX 77030-4009. Phone: 713-563-9184; Fax: 713-792-0309; E-mail: iiwistuba@mdanderson.org or Matthew Meyerson. E-mail: Matthew_Meyerson@dfci.harvard.edu.

©2008 American Association for Cancer Research.

doi:10.1158/1940-6207.CAPR-08-0177

hybridization (CGH) profiling has been previously applied to oral squamous cell carcinoma (16–18), leading to the identification of novel chromosomal regions (e.g., 5p15.2 and 11q.22.2-22.3) frequently altered in this neoplasm (16). Earlier studies using GCH and single-nucleotide polymorphism (SNP) array have already shown that chromosomal abnormalities increase with the progression of BE lesions, but they analyzed only a few tissue specimens (19, 20). Both of the present studies showed that microarray-based DNA copy number analyses can reveal the progression of chromosomal abnormalities that parallels the clinical and histopathologic progression of premalignancy. The findings of these studies also have some potential to be used to predict the risk of progression of either oral leukoplakia or BE.

Tsui et al. (14) used CGH to evaluate genetic alterations on chromosome 3p in 47 oral premalignant lesions with high-grade dysplastic features (severe dysplasia and carcinoma *in situ*) and compared these alterations with findings in 23 oral squamous cell carcinomas; all samples were formalin-fixed and paraffin-embedded tissues. This study was stimulated by previous laboratory work of this group (21, 22) and others (2, 10, 11), showing that chromosome 3p is frequently altered in the early pathogenesis of oral carcinoma and has been associated with the risk of premalignant lesion progression (2, 10, 11, 21, 22). High-grade dysplastic lesions exhibited six recurrent regions of losses on 3p including 3p25.3-p26.1, 3p25.1-p25.3, 3p24.1, 3p21.31-p22.3, 3p14.2, and 3p14.1, which overlapped losses found in invasive carcinoma. Next, the authors examined these regions in 24 low-grade dysplastic lesions with known outcomes, including 2 hyperplasias and 22 mild and moderate dysplasias, and determined that 3p losses were identified more frequently in low-grade dysplastic lesions with progressive behavior (78%) compared with nonprogressing lesions (20%). This interesting observation obviously will require verification in a larger cohort of patients. In our opinion, one remarkable finding of this report is that the size of 3p segmental losses, or discontinuous LOH (alternating segments showing loss and retention; portrayed in Fig. 1), increased with histologic stage; segmental 3p losses were detected in premalignant lesions, whereas whole-arm loss occurred mainly in invasive tumors.

Li et al. (15) used a medium-density SNP array (containing ~33,000 SNPs) to investigate genome-wide chromosomal copy number changes in whole frozen tissue specimens (obtained by endoscopy or surgical resection) of multiple stages of BE and esophageal adenocarcinoma from 42 patients. This investigation is from the same laboratory that established the working model for BE molecular progression (3) and one of the first laboratories to use SNP arrays for the analysis of human premalignancy (23). The mucosal esophageal tissues used in this cross-sectional study included 24 early-stage BE specimens (with or without aneuploidy and that did not develop cancer during follow-up), 10 late-stage BE specimens (with microscopic invasion), and 8 grossly invasive esophageal adenocarcinoma specimens. Genome-wide copy losses and LOH (the most frequent changes) and copy gains increased in frequency and size between early and late BE, with SNP abnormalities increasing from 2% in early stages to >30% in late stages. A set of statistically significant events was unique to either early or late BE stage, and few chromosomal regions with changes were common in all stages of progression. It is interesting that the total number of genome-wide SNP alterations (gains, losses, and LOH) was highly correlated with DNA content aneuploidy and was sensitive and specific for identifying patients with concurrent esophageal adenocarcinoma. As with chromosome 3p in the study of Tsui et al. (14), Li et al. found that the sizes of chromosome abnormalities were small in early-stage BE compared with late stages, including invasive adenocarcinoma, except in the case of 9p LOH.

How do the articles by Tsui et al. (14) and Li et al. (15) contribute to the general concepts of the parallel progressions of genetic and clinical/histopathologic changes in premalignancy (described above)? First, these investigators have successfully applied high-throughput DNA chromosomal abnormality analysis technologies, array CGH and SNP array analysis, to small specimens of premalignant lesion tissue, formalin-fixed and paraffin-embedded specimens (CGH), and frozen samples (SNP). We hope that these two reports will raise enthusiasm for similar new studies in these and other neoplasm models.

Second, these studies show that chromosomal abnormalities (mostly deletions) at the 3p chromosome (14) and genome-wide (15) levels commence early in oral and BE premalignancies and progress with the increasing severity of epithelial changes and that these chromosomal changes follow a sequence defining early and late molecular changes. Third, these two studies have suggested that the size of chromosomal deletions (at 3p or genome-wide) increases with the severity of histopathologic changes. This interesting finding has been reported previously in the pathogenesis of other epithelial tumors, such as lung tumors, where it has been observed in LOH studies using PCR-based amplification of multiple microsatellites in precisely microdissected histologically normal and premalignant epithelia obtained from lung cancer patients and high-risk individuals (smokers; refs. 24–26). LOH at multiple chromosome 3p (24, 26) and 8p (25) sites was shown to commence at the stage of normal epithelium and to increase with progressive histologic changes in the lung squamous cell carcinoma histology progression model (similar in ways to oral squamous carcinoma). This study used 24 microsatellite markers spanning six continuous chromosome 3p regions in showing that deletions in 3p progressed from dysplasia to invasive squamous carcinoma (Fig. 1). In invasive tumors and carcinoma *in situ*, most of the 3p arm was deleted and the extent of the deletions was greater in all cases than that in corresponding normal and premalignant foci (24). These findings were subsequently expanded in a LOH analysis (using 54 microsatellite markers on 3p) that included samples from a wider spectrum of premalignancy and lung cancer specimens, including samples from smokers and cancers with different histologies (nonsmall and small cell). This detailed allelotyping analysis identified multiple areas of discontinuous LOH and thus multiple breakpoints throughout the 3p arm in many tumor and bronchial epithelial samples (26). Allelic losses present in lung premalignant lesions were not random and followed a sequence, with the earliest and most frequent allelic loss occurring at the 3p21.3 region (24, 26).

Although the two articles in this issue of the journal report that discontinuous LOH marks progression of premalignancy, this LOH also could be an artifact of impurity within the premalignant lesions. In samples with an admixture of DNA from malignant and normal cells, some chromosomal segments might seem to be altered and others not altered, although the degree of alteration could be the same across the entire chromosome arm because of nonuniform behavior of copy number probes, which fail to reach a threshold level for detecting an alteration. This failure has been shown for SNP array analysis of mixtures of tumor cell line DNA with matched normal DNA; increasing admixture of normal DNA led to the appearance of discontinuous LOH because of incomplete detection involving regions of uninterrupted chromosome 3p LOH in the pure tumor (27). Further improvements in genomics technologies coupled with careful histologic analysis will be required to address this question.

The mechanism for chromosomal deletions in individual premalignant samples is not well established. LOH is considered to occur through the loss of a whole chromosome because of inappropriate chromosomal segregation at mitosis and also through genetic alterations that change chromosomal structures (28). Whole, terminal, and interstitial chromosome physical deletions have been shown to cause LOH on several chromosomal arms in human tumors (29). On the other hand, mitotic recombination and gene conversion seem to be additional mechanisms causing LOH (30, 31). Unbalanced translocations, which have been confirmed by cytogenetic studies in certain human tumors, have been implicated in LOH development (32). In summary, six chromosomal aberrations (whole chromosome, terminal and interstitial deletions, mitotic recombination, gene conversion, and unbalanced translocation) are considered to be responsible for LOH in human carcinogenesis (Fig. 2). However, the contribution of each chromosome alteration to the occurrence of LOH has been examined only for a few chromosomal regions in few tumor types. It was determined recently that 80% of LOHs are partial chromosome deletions

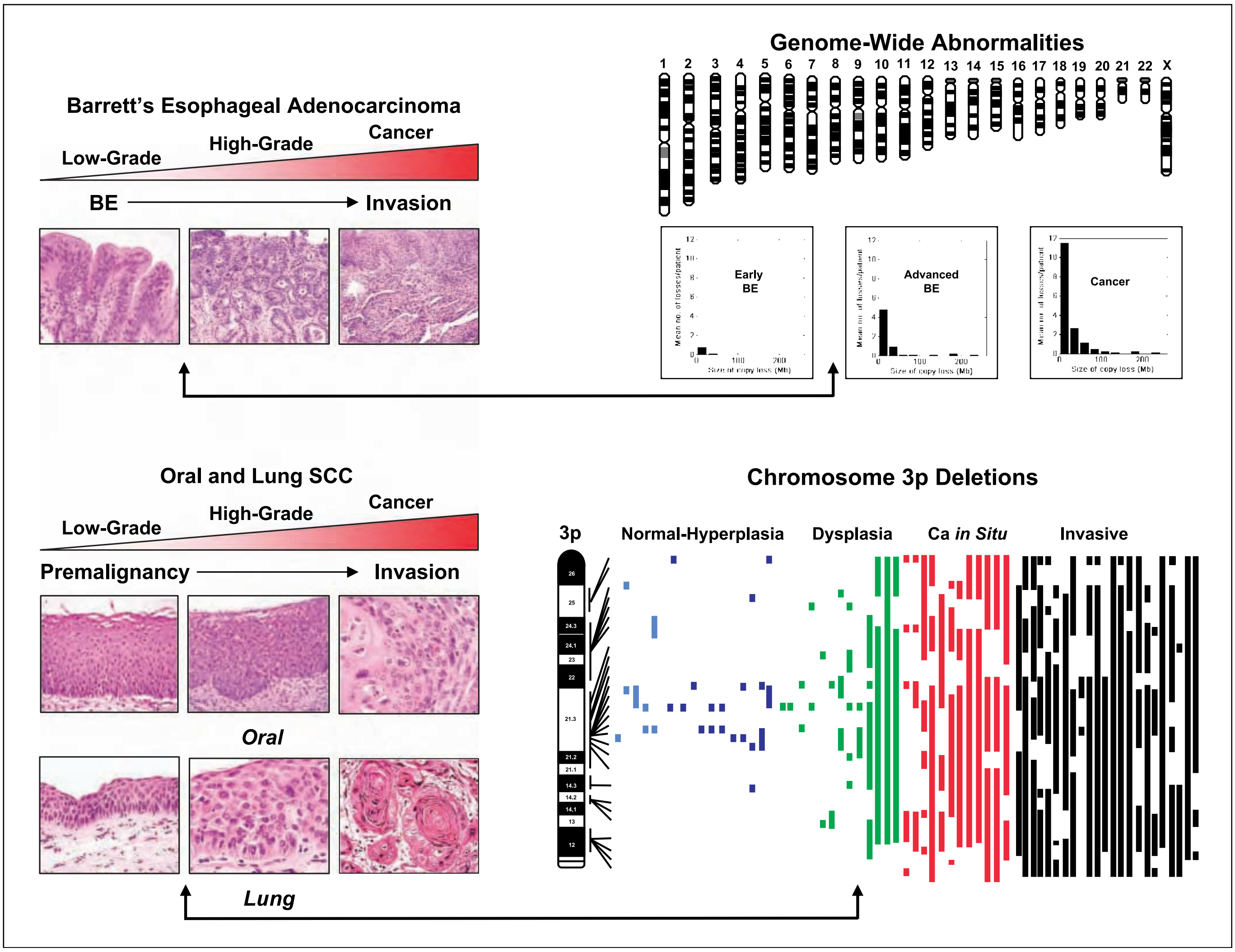

Barrett's Esophageal Adenocarcinoma
Low-Grade
High-Grade
Cancer
BE
Invasion
Genome-Wide Abnormalities
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
Early BE
Advanced BE
Cancer
Mean no. of losses/patient
Size of copy loss (Mb)
0 100 200
0 2 4 6 8 10 12
Oral and Lung SCC
Low-Grade
High-Grade
Cancer
Premalignancy
Invasion
Oral
Lung
Chromosome 3p Deletions
3p
Normal-Hyperplasia
Dysplasia
Ca in Situ
Invasive
26
25
24.3
24.1
23
22
21.3
21.2
21.1
14.3
14.2
14.1
13
12

Fig. 1. Progression of premalignancy and molecular changes in Barrett's esophageal adenocarcinoma (*top*) and oral and lung squamous cell carcinomas (*SSC*; *bottom*). Chromosomal deletions at genome-wide and the 3p chromosome levels begin early in esophageal premalignancy (or BE) and, in oral premalignancy, progress and increase in size with increasing histopathologic severity (14, 15). Li et al. (15) showed that the number of larger copy losses (Mb) was significantly higher in advanced BE and esophageal adenocarcinoma compared with early BE lesions (*top*). In oral premalignancy progression, Tsui et al. (14) showed that the size of 3p segmental losses increased with histologic stage, as has been shown previously in progression to lung squamous cell carcinoma (*bottom*; refs. 24–26). In the lung squamous cell carcinoma model, discontinuous segmental losses, or LOH, at 3p are detected in normal epithelium, hyperplasia, and some dysplasias, whereas the whole arm is lost in invasive and *in situ* carcinomas and in a subset of dysplastic lesions; the vertical colored bars (*bottom right*) indicate the size of deletions; the gaps, or retentions, between the bars indicate discontinuous LOH. Histology pictures of BE progression are courtesy of Elizabeth Montgomery, M.D., and of oral leukoplakia progression are courtesy of Adel El-Naggar, M.D., Ph.D.

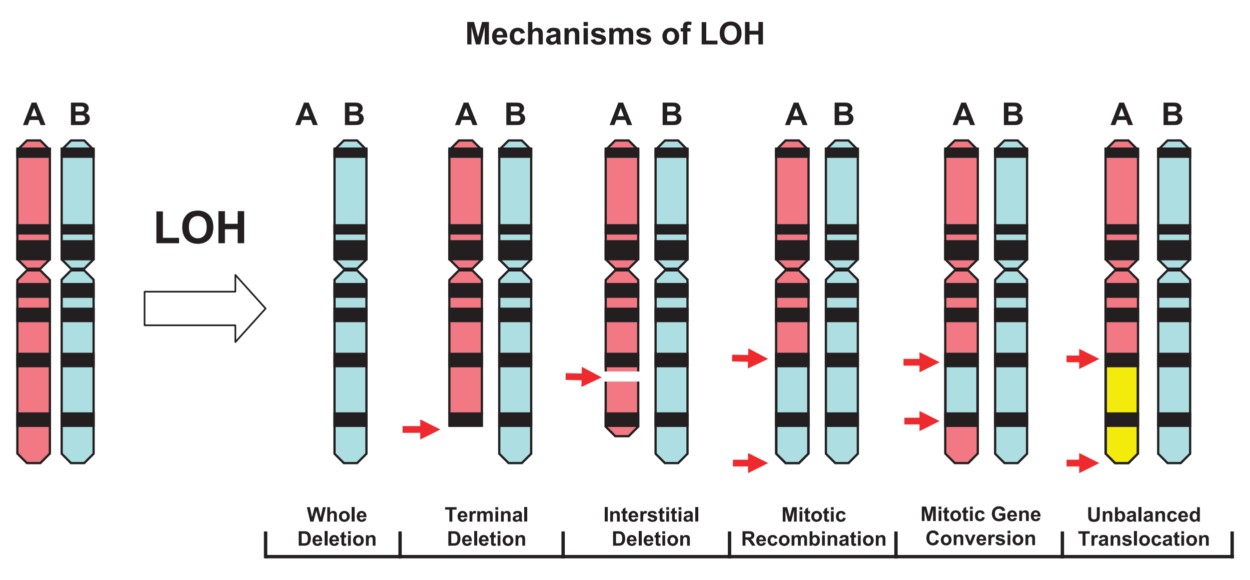


Fig. 2. Chromosomal alterations considered to be responsible for LOH in human carcinogenesis. The whole-chromosome, terminal, and interstitial deletions and unbalanced translocation represent copy number changes. The red arrows indicate sites of chromosomal breaks.

(involving the several chromosomal alterations mentioned above), whereas the remaining 20% were whole-chromosome deletions (33); these results came from integrating information on breakpoints for DNA copy number changes (obtained by a CGH) with numerical and structural chromosomal alterations (obtained by spectral karyotyping in lung cancer cell lines). We know that DNA copy number analyses of tumors have limitations due to the phenomenon that most of these alterations in invasive cancers are large, spanning many genes (tens to hundreds), including many that likely are not involved in oncogenesis. Therefore, studies of premalignant lesions with known outcomes are necessary to better define chromosomal loci and genes responsible for tumor development in humans.

The reports of Tsui et al. (14) and Li et al. (15) on the tumor models oral squamous cell carcinoma and Barrett's esophageal adenocarcinoma give new insights into and hope for understanding the highly complex nature of the progression of premalignancy to cancer. Despite their different pathogenetic features, these two tumors shared some common genetic progression characteristics. The application of high-throughput DNA technology in these models showed the promise of this approach and the merit of extending these methods to other premalignant diseases. It also showed the potential of CGH and SNP arrays to identify new candidate biomarkers and measures of clonal diversity, both of which can be used in patient management and assessment of cancer risk. Is less more? High-throughput approaches generate increasing amounts of data from smaller and smaller tissue samples, and chromosomal deletions appear to increase progression of premalignancy. These examples of "less is more" represent important advances in our understanding of the molecular nature of carcinogenesis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Mao L, Hong WK, Papadimitrakopoulou VA. Focus on head and neck cancer. Cancer Cell 2004;5:311–6.

2. Paulson TG, Reid BJ. Focus on Barrett's esophagus and esophageal adenocarcinoma. Cancer Cell 2004;6:11–6.

3. Wistuba I, Gazdar A. Lung cancer preneoplasia. Annu Rev Pathol Mech Dis 2006;1:331–48.

4. Wistuba II, Mao L, Gazdar AF. Smoking molecular damage in bronchial epithelium. Oncogene 2002;21:7298–306.

5. Minna JD, Gazdar A. Focus on lung cancer. Cancer Cell 2002;1:49–52.

6. De Marzo AM, Meeker AK, Zha S, et al. Human prostate cancer precursors and pathobiology. Urology 2003;62:55–62.

7. Dinney CP, McConkey DJ, Millikan RE, et al. Focus on bladder cancer. Cancer Cell 2004;6:111–6.

8. Knudson AG. Hereditary cancers: clues to mechanisms of carcinogenesis. Br J Cancer 1989;59:661–6.

9. Croce CM. Oncogenes and cancer. N Engl J Med 2008;358:502–11.

10. Papadimitrakopoulou VA, Clayman GL, Shin DM, et al. Biochemoprevention for dysplastic lesions of the upper aerodigestive tract. Arch Otolaryngol Head Neck Surg 1999;125:1083–9.

11. Mao L, Lee JS, Fan YH, et al. Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. Nat Med 1996;2:682–5.

12. Maley CC, Reid BJ. Natural selection in neoplastic progression of Barrett's esophagus. Semin Cancer Biol 2005;15:474–83.

13. Barrett MT, Sanchez CA, Prevo LJ, et al. Evolution of neoplastic cell lineages in Barrett oesophagus. Nat Genet 1999;22:106–9.

14. Tsui IF, Rosin MP, Zhang L, Ng RT, Lam WL. Frequent genetic alterations occurring in oral premalignant lesions: analysis of chromosome 3p. Cancer Prev Res 2008;1:424–9.

15. Li X, Galipeau PC, Sanchez CA, et al. SNP-based genome-wide chromosome copy change, LOH, aneuploidy in Barret's esophagus neoplastic progression. Cancer Prev Res 2008;1:413–23.

16. Baldwin C, Garnis C, Zhang L, Rosin MP, Lam WL. Multiple microalterations detected at high frequency in oral cancer. Cancer Res 2008;1:413–24.

17. Liu CJ, Lin SC, Chen YJ, Chang KM, Chang KW. Array-comparative genomic hybridization to detect genomewide changes in microdissected primary and metastatic oral squamous cell carcinomas. Mol Carcinog 2006;45:721–31.

18. Nakaya K, Yamagata HD, Arita N, et al. Identification of homozygous deletions of tumor suppressor gene FAT in oral cancer using CGH-array. Oncogene 2007;26:5300–8.

19. Walch AK, Zitzelsberger HF, Bruch J, et al. Chromosomal imbalances in Barrett's adenocarcinoma and the metaplasia-dysplasia-carcinoma sequence. Am J Pathol 2000;156:555–66.

20. Lai LA, Paulson TG, Li X, et al. Increasing genomic instability during premalignant neoplastic progression revealed through high resolution array-CGH. Genes Chromosomes Cancer 2007; 46:532–42.

21. Ng IO, Xiao L, Lam KY, Yuen PW, Ng M. Microsatellite alterations in squamous cell carcinoma of the head and neck—clustering of loss of heterozygosity in a distinct subset. Oral Oncol 2000;36: 484–90.

22. Rosin MP, Cheng X, Poh C, et al. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. Clin Cancer Res 2000;6: 357–62.

23. Mei R, Galipeau PC, Prass C, et al. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res 2000; 10:1126–37.

24. Wistuba II, Behrens C, Milchgrub S, et al. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. Oncogene 1999;18:643–50.

25. Wistuba II, Behrens C, Virmani AK, et al. Allelic losses at chromosome 8p21-23 are early and frequent events in the pathogenesis of lung cancer. Cancer Res 1999;59:1973–9.

26. Wistuba II, Behrens C, Virmani AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.

27. Lindblad-Toh K, Tanenbaum DM, Daly MJ, et al. Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol 2000;18:1001–5.

28. Weinberg RA. Tumor suppressor gene. In: Weinberg RA, editor. The biology of cancer. New York: Garland Science, Taylor and Francis Group LLC; 2007. p. 209–54.

29. Lasko D, Cavenee W, Nordenskjold M. Loss of constitutional heterozygosity in human cancer. Annu Rev Genet 1991;25:281–314.

30. Hagstrom SA, Dryja TP. Mitotic recombination map of 13cen-13q14 derived from an investigation of loss of heterozygosity in retinoblastomas. Proc Natl Acad Sci U S A 1999;96:2952–7.

31. Adams J, Williams SV, Aveyard JS, Knowles MA. Loss of heterozygosity analysis and DNA copy number measurement on 8p in bladder cancer reveals two mechanisms of allelic loss. Cancer Res 2005;65:66–75.

32. Roschke AV, Tonon G, Gehlhaus KS, et al. Karyotypic complexity of the NCI-60 drug-screening panel. Cancer Res 2003;63:8634–47.

33. Ogiwara H, Kohno T, Nakanishi H, Nagayama K, Sato M, Yokota J. Unbalanced translocation, a major chromosome alteration causing loss of heterozygosity in human lung cancer. Oncogene 2008;27: 4788–97.

EGFR-T790M Is a Rare Lung Cancer Susceptibility Allele with Enhanced Kinase Activity

Haris Vikis,[1] Mitsuo Sato,[2] Michael James,[1] Daolong Wang,[1] Yian Wang,[1] Min Wang,[1] Dongmei Jia,[1] Yan Liu,[1] Joan E. Bailey-Wilson,[3] Christopher I. Amos,[4] Susan M. Pinney,[5] Gloria M. Petersen,[6] Mariza de Andrade,[6] Ping Yang,[6] Jonathan S. Wiest,[7] Pamela R. Fain,[8] Ann G. Schwartz,[9] Adi Gazdar,[2] Colette Gaba,[10] Henry Rothschild,[11] Diptasri Mandal,[11] Elena Kupert,[5] Daniela Seminara,[7] Avinash Viswanathan,[1] Ramaswamy Govindan,[1] John Minna,[2] Marshall W. Anderson,[5] and Ming You[1]

[1]Washington University, St. Louis, Missouri; [2]University of Texas Southwestern Medical Center, Dallas, Texas; [3]National Human Genome Research Institute, Bethesda, Maryland; [4]M. D. Anderson Cancer Center, Houston, Texas; [5]University of Cincinnati, Cincinnati, Ohio; [6]Mayo Clinic College of Medicine, Rochester, Minnesota; [7]National Cancer Institute, Rockville, Maryland; [8]University of Colorado, Denver, Colorado; [9]Karmanos Cancer Institute, Detroit, Michigan; [10]Medical University of Ohio, Toledo, Ohio; and [11]Louisiana State University Health Science Center, New Orleans, Louisiana

Abstract

The use of tyrosine kinase inhibitors (TKI) has yielded great success in treatment of lung adenocarcinomas. However, patients who develop resistance to TKI treatment often acquire a somatic resistance mutation (T790M) located in the catalytic cleft of the epidermal growth factor receptor (EGFR) enzyme. Recently, a report describing EGFR-T790M as a germ-line mutation suggested that this mutation may be associated with inherited susceptibility to lung cancer. Contrary to previous reports, our analysis indicates that the T790M mutation confers increased Y992 and Y1068 phosphorylation levels. In a human bronchial epithelial cell line, overexpression of EGFR-T790M displayed a growth advantage over wild-type (WT) EGFR. We also screened 237 lung cancer family probands, in addition to 45 bronchoalveolar tumors, and found that none of them contained the EGFR-T790M mutation. Our observations show that EGFR-T790M provides a proliferative advantage with respect to WT EGFR and suggest that the enhanced kinase activity of this mutant is the basis for rare cases of inherited susceptibility to lung cancer. [Cancer Res 2007;67(10):4665–70]

Introduction

Recent work has identified a series of somatic mutations in exons 18 to 21 of epidermal growth factor (EGF) receptor (*EGFR*) that render lung tumors responsive to the gefitinib and erlotinib therapeutics (1–3). However, in patients that progress after drug treatment, it has been observed that a secondary "resistance" mutation is often acquired in exon 20 (4–6). This mutation, T790M, arises somatically in ~50% of these cases (4–6). In chronic myelogenous leukemia (CML) patients, an estimated 50% to 90% of tumors with acquired resistance have the analogous resistance mutation (T315I) in BCR-ABL (7). Interestingly, Bell et al. (8) report the first identification of the T790M mutation in the germ line of a European family that developed lung adenocarcinoma with bronchoalveolar (BAC) differentiation. This observation suggests that this mutation may be associated with genetic susceptibility to lung cancer and may underlie familial predisposition to the disease. The allele, although common in drug-treated tumors, seems extremely rare in the germ line of the general population, as the authors report no mutation observed in 782 alleles sequenced. Many groups have reported that the kinase activity of the EGFR-T790M–resistant mutant is indistinguishable from wild-type (WT) EGFR (4–6, 9, 10). What is somewhat perplexing is how T790M would confer susceptibility if its activity were identical to the WT molecule. Our analysis shows that T790M in fact does exhibit enhanced autophosphorylation at Y992 and Y1068, and this mutation is associated with a proliferative advantage in a human bronchial epithelial cell line. Interestingly, the mutation seems to be rare as it was not found in any of the familial or sporadic lung cancer populations we screened.

Materials and Methods

Cell lines. HEK293T, HEK293FT, and COS-7 cells were maintained in DMEM supplemented with 10% fetal bovine serum. Immortalized human bronchial epithelial cells (HBEC3-KT) were maintained in keratinocyte serum-free medium (with 50 μg/mL bovine pituitary extract and 5 ng/mL EGF). All cells were grown at 37°C in a humidified incubator with 5% CO_2.

Plasmids, transfection, and viral infection. Mammalian expression plasmids encoding for human EGFR, EGFR-T790M, and EGFR-L858R were kind gifts from William Pao (Memorial Sloan-Kettering Cancer Center, New York, NY). HEK293T and COS-7 cells were transfected with the indicated EGFR plasmids using LipofectAMINE 2000 (Invitrogen). Gefitinib was obtained from Chemoprevention Branch, National Cancer Institute (Bethesda, MD) and added to cells 10 h before lysis.

To introduce WT and mutant EGFRs into HBEC3 cells, we used the pLenti6/V5 Directional TOPO Cloning kit (Invitrogen). Construction of pLenti-wt-EGFR was described previously (11). The T790M and L747_E749 deletion mutations were introduced into pLenti-wt-EGFR by using site-directed mutagenesis (Stratagene). pLenti-KRASV12 vector was constructed by cloning KRASV12 fragment from pBabe-KRASV12-hyg (11) into pLenti6/V5 vector. pLenti6/V5-GW/*lacZ* (Invitrogen) was used as a control. Viral transduction of HBEC3 cells was done following the manufacturer's instructions. Briefly, the 293FT cells were transiently transfected with viral vector together with ViraPower (Invitrogen). Forty-eight hours after transfection, the supernatant of the 293FT cells was harvested and passed through a 0.45-μm filter and frozen at −80°C. The supernatant was used for infection after addition of 4 μg/mL polybrene (Sigma). Forty-eight hours after infection, drug selection of infected cells was started with 5 μg/mL blasticidin (Invitrogen) and continued for 7 days.

Note: H. Vikis, M. Sato, M. James, and D. Wang contributed equally to this work.

Requests for reprints: Ming You, Department of Surgery and The Alvin J. Siteman Cancer Center, Washington University, 660 Euclid Avenue, Box 8109, St. Louis, MO 63110. Phone: 314-362-9294; Fax: 314-362-9366; E-mail: youm@wustl.edu.

©2007 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-0217

Colony formation assays. Liquid colony formation assays were done as described previously (11). Briefly, 200 viable cells were plated in triplicate on 100-mm plates and cultured in keratinocyte serum-free medium supplemented with 50 μg/mL of bovine pituitary extract without EGF. Surviving colonies were counted 10 days later after staining with methylene blue, and colonies >3 mm in diameter were counted.

Western blot and antibodies. Cell lysates were prepared in LDS sample buffer (Invitrogen) and electrophoresed on NuPAGE gels (Invitrogen). Protein was transferred to polyvinylidene difluoride membranes and blotted using the antibodies as indicated: anti-EGFR (Cell Signaling Technology), anti–EGFR-Y1068 (Cell Signaling Technology), anti–EGFR-Y992 (Cell Signaling Technology), anti–mitogen-activated protein kinase (MAPK; Cell Signaling Technology), anti–phosphorylated MAPK (Thr^{202}/Tyr^{204}; Cell Signaling Technology), anti–phosphorylated AKT (Ser^{473}; Cell Signaling Technology), anti-AKT (Cell Signaling Technology), anti–α-tubulin (Santa Cruz Biotechnology), and anti–cyclin D1 (Santa Cruz Biotechnology).

***EGFR* exon 20 genotyping.** Amplification of human EGFR exon 20 was done via standard PCR methods using forward (5′-GACACTGACG-TGCCTCTCC-3′) and reverse (5′-TTATCTCCCCTCCCCGTATC-3′) primers. PCR products were electrophoresed on agarose gels, purified, and subjected to standard DNA sequencing. The EGFR-T790M mutation is deduced by the genotype (C/T) at position 86 in exon 20.

Lung cancer DNAs. The Genetic Epidemiology of Lung Cancer Consortium (GELCC) has accrued over 700 families with three or more first-degree relatives with lung cancer (12), of which this study genotyped 237 individual probands. Paraffin blocks of lung tumors (45) with BAC differentiation (and 32 corresponding normals), in addition to the fresh frozen lung tumor DNAs, were obtained from the Washington University Tissue Procurement Center as paraffin blocks. For the blocks, genomic DNA was prepared by slicing several paraffin curls and adding xylenes to dissolve the paraffin. The remaining tissue was washed with ethanol, dried, and resuspended in PCR buffer (0.5% Tween 20, 0.5% NP40) and digested in proteinase K (1 mg/mL) at 55°C overnight. Chloroform/isoamyl alcohol (24:1) was added to the digest and soluble DNA was isolated from the aqueous phase.

Results

To address previous observations that EGFR-T790M activity is the same as the WT molecule, we did transient transfections of EGFR and EGFR-T790M expression constructs in COS-7 and 293T cells. Overexpression of EGFR and the T790M mutant in these cell lines clearly indicated that the phosphorylation status at Y1068 is enhanced with respect to WT (Fig. 1*A–E*). Tyr^{1068} is an EGFR autophosphorylation site that couples receptor activation to Ras signaling (13). Interestingly, our observations are in disagreement with previously published reports (4–6, 9, 10). Our data also show that phosphorylation levels between WT and T790M are indistinguishable at higher expression levels (Fig. 1*C*), yet differences are more apparent at lower expression levels. As a control, we also showed the drug-resistant nature of EGFR-T790M in comparison with WT EGFR (Fig. 1*D*). We further addressed the activation status of MAPK, a downstream molecule phosphorylated and activated by EGFR. We observed that phosphorylated MAPK

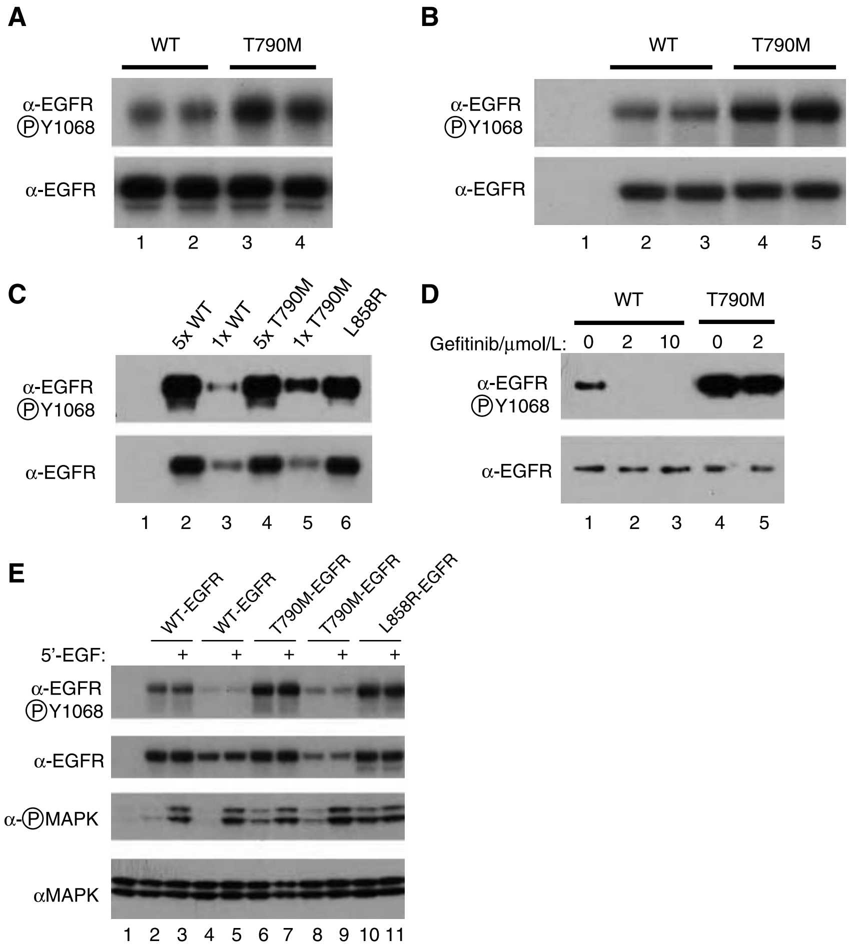


Figure 1. Phosphorylation of Y1068 is enhanced in EGFR-T790M. HEK293 cells (*A*) and COS-7 cells (*B*) were transfected with the indicated DNAs and lysed 48 h after transfection. Lysates were immunoblotted with anti-EGFR and anti–EGFR-Y1068 as indicated. Phosphorylation of Y1068 is enhanced in the EGFR-T790M mutant. *C,* phosphorylation and activation of EGFR-T790M are evident at lower levels of protein expression compared with higher levels in COS-7 cells (*lane 3* versus *lane 5*). *D,* gefitinib does not affect the phosphorylation status of T790M-EGFR. *E,* phosphorylation of MAPK is enhanced in EGFR-T790M–expressing (versus EGFR-WT expressing) cells. Cells were stimulated with EGF (10 ng/mL) for 5 min.

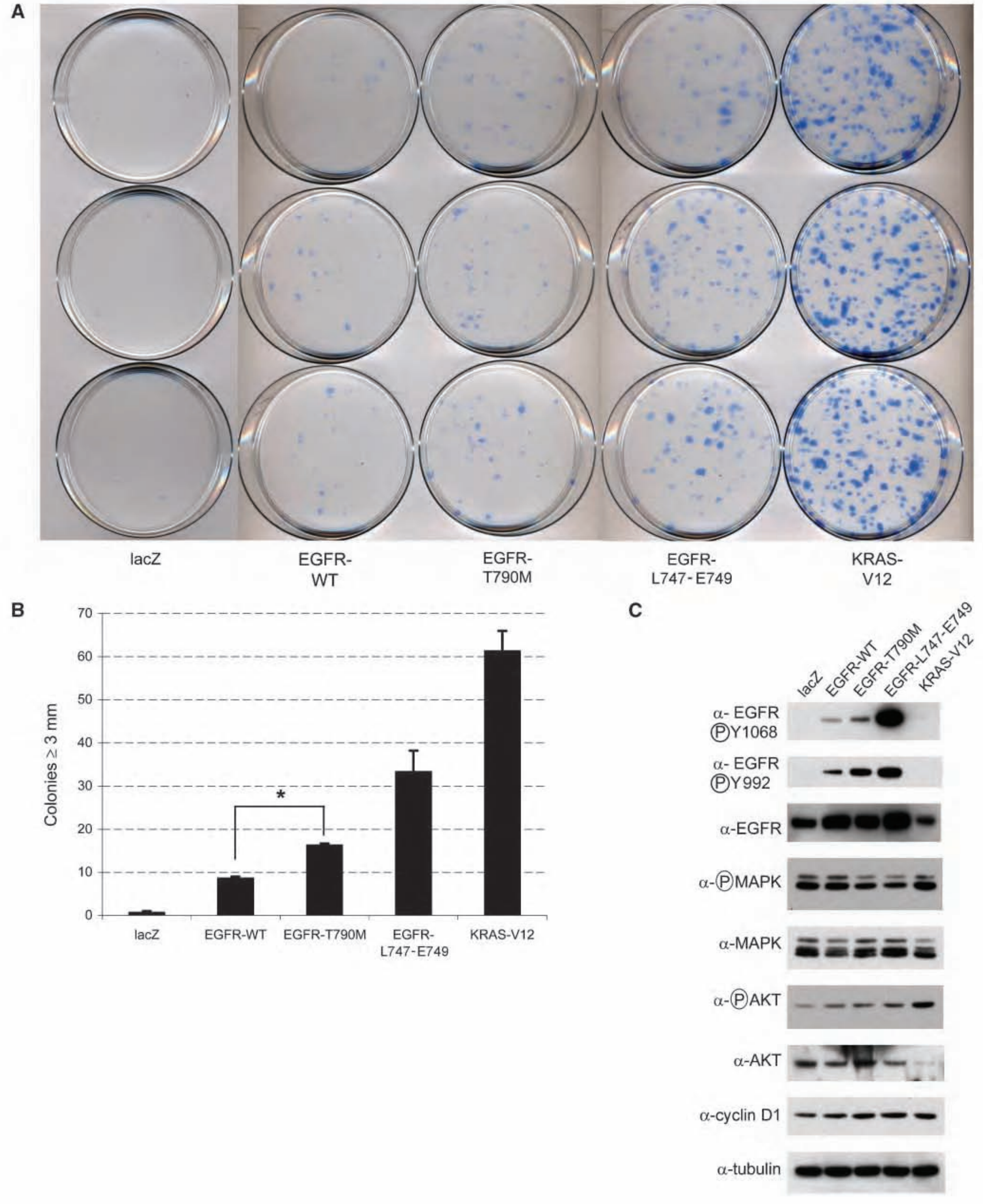
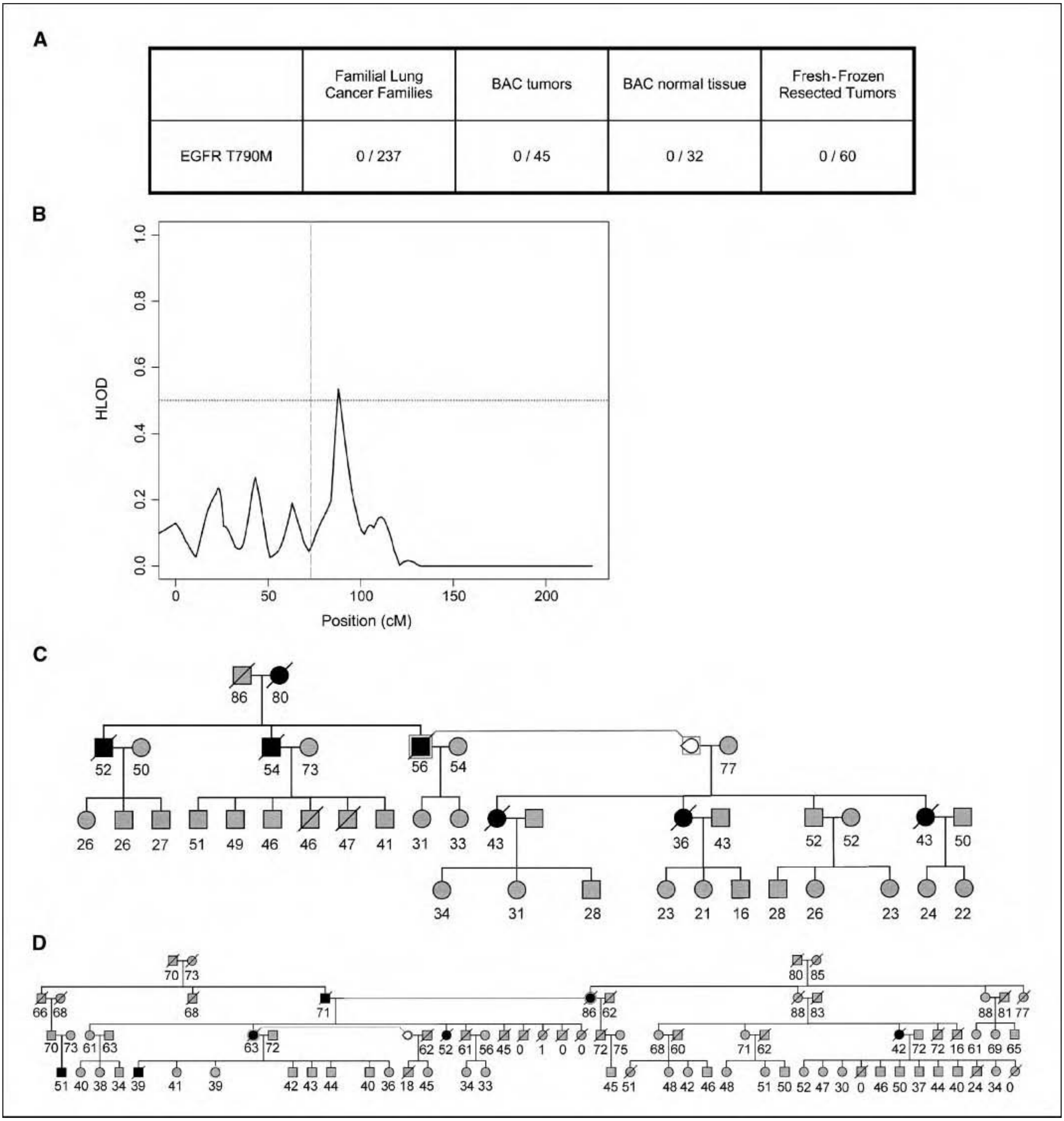


Figure 2. Colony formation growth assay for EGFR-T790M. Viral transduction and selection of HBEC cells stably expressing EGFR were done as described previously (11). *A,* cells (200) were plated in triplicate and grown in keratinocyte serum-free medium (with 50 μg/mL bovine pituitary extract and without EGF). After 10 d, cells were stained with methylene blue. *B,* stained cell colonies >3 mm in diameter were counted. *, $P < 0.0001$, two-tailed *t* test. *C,* Western blots for EGFR, EGFR-Y1068, EGFR-Y992, MAPK, MAPK (T202/Y204), AKT, AKT (S473), cyclin D1, and α-tubulin in the HBEC stable cell lines were done as indicated.

(Thr^{202}/Tyr^{204}) levels were enhanced in the T790M mutant compared with WT EGFR (Fig. 1*E*).

We also tested whether EGFR-T790M expression in an immortalized human bronchial epithelial cell (HBEC3) line affected growth properties of these cells. Cells were transduced by lentivirus expressing EGFR and EGFR mutants and selected for stable integration as done previously (11). Two hundred cells were seeded on plates without EGF supplementation and allowed to grow for



	Familial Lung Cancer Families	BAC tumors	BAC normal tissue	Fresh-Frozen Resected Tumors
EGFR T790M	0 / 237	0 / 45	0 / 32	0 / 60

Figure 3. EGFR genotyping and linkage analysis for chromosome 7 with SimWalk2. *A,* the EGFR-T790M mutation was not observed in familial lung cancer blood DNA, BAC tumors and normals, or fresh-frozen resected tumors. *B,* the genetic position for EGFR is at ~73 cM (based on interpolation), between markers D7S1818 (70 cM) and D7S3046 (79 cM). Simwalk2 analysis was based on the parametric method using the genetic model reported by Moscatello et al. (18). Fifty-five pedigrees with five or more affected members were used. *C* and *D,* two representative pedigrees of GELCC collected families with predisposition to lung cancer. In (*C*) and (*D*), ● (females) or ■ (males) represent lung, throat, or laryngeal cancer. *Numbers below each individuals,* sample numbers; *numbers in brackets,* individual ages at the time the pedigree was constructed; *slashes,* dead individuals.

10 days. We observed that the EGFR-T790M–expressing cells had a greater number of large (≥3 mm in diameter) colonies compared with WT EGFR but less than the activated L747-E749 deletion (Fig. 2*A* and *B*). These results indicate that EGFR-T790M does in fact provide a growth advantage over WT EGFR. When we analyzed the status of EGFR phosphorylation at Y992 and Y1068 in these cells, we saw an increase in the EGFR-T790M mutant (Fig. 2*C*), consistent with what was observed in HEK293T and COS-7 cells (Fig. 1*A–D*).

To address possible downstream signaling mechanisms that are affected by mutant EGFR, we looked at the phosphorylation/activation status of MAPK and AKT (Fig. 2*C*). In contrast to our observations in 293T cells, we did not see MAPK activation induced by T790M. However, this is consistent with previously published results where no obvious changes are seen with the deletion mutant, yet colony formation is enhanced (11). Transcriptional changes, such as up-regulation of cyclin D1 levels, were recently reported to be a result of EGFR mutant signaling (14); however, we observed no changes induced by the mutants (Fig. 2*C*).

If *EGFR-T790M* is a human susceptibility allele, it is necessary to determine its prevalence in families with high susceptibility to lung cancer. With these families, a set of 52 extended pedigrees was used to show linkage on chromosome 6q23-25 (12). A concerted effort by the GELCC has attempted to identify the gene(s) associated with this susceptibility. We proposed that if the T790M mutation is a common variant for predisposition to lung cancer, it could potentially be enriched in probands from lung cancer families. Genomic DNA from 237 probands representing lung cancer families with more than three affected individuals was analyzed. PCR amplification and DNA sequencing of the resistance mutation in exon 20 were done. Interestingly, we did not observe the T790M mutation in any of the family probands analyzed or in any of 60 random fresh-frozen resected lung tumors (Fig. 3*A*). This would suggest that the T790M mutation is not enriched in our population and is likely a minor contributor to genetic susceptibility in familial lung cancers. Furthermore, in the aforementioned GELCC linkage study, the analysis of 52 families (and family subsets within) did not reveal a significant logarithm of odds score on or near 7p11, where the *EGFR* gene is located, which also suggests that the mutation is not a major contributor to familial predisposition to lung cancer (Fig. 3*B*; ref. 12). Representative families used in this study are indicated in Fig. 3*C* and *D*. We further tested to see if T790M might in fact be solely responsible for predisposition to adenocarcinomas with BAC differentiation. We sequenced 45 BAC tumors and 32 of the corresponding normal tissues and did not observe the T790M mutation in any of these samples (Fig. 3*A*). This further suggests the rarity of this potential predisposing mutation.

Discussion

Several lines of evidence exist that lend support to T790M as a putative susceptibility allele. T790M has been detected somatically in tumors that have not been treated with gefitinib or erlotinib, which may suggest that this mutant may have arisen during drug-free cancer progression (15). Furthermore, the NCI-H1975 BAC cell line, which has never undergone tyrosine kinase inhibitor treatment, has both activating L858R and resistant T790M mutations, suggestive that T790M may be growth promoting (5). Bell et al. (8) also have observed that the T790M mutation seems to occur in *cis* with the activating mutations. Perhaps more persuasive is that the analogous resistance mutation in CML patients, BCR-ABL-T315I, displays increased *in vitro* kinase activity (16). Similarly, the analogous mutations in Src (T341M) and FGFR1 (V561M) also result in increased phosphorylation and activation (10). Why T790M in EGFR does not reportedly function in a similar manner is unclear.

Our results suggest that the T790M mutation may in fact provide a proliferative advantage in normal cells by increasing kinase activity and downstream signaling. Our data suggest that EGFR-T790M does in fact exhibit higher kinase activity than WT molecule in HEK293T and COS-7 cell lines. Overexpression in our HBEC3 cell line showed increased tyrosine phosphorylation at Y992 and Y1068 and increased colony formation. Increased proliferation is most evident in the absence of EGF in the medium, suggesting that the activating T790M EGFR mutation is involved in the proliferative effects observed. EGFR kinase activity is essential for oncogenic transformation. EGFR extracellular domain deletion mutants and overexpression of EGFR are commonly found in human cancers (17, 18). Our data also show that phosphorylation levels between WT and T790M are indistinguishable at higher expression levels (Fig. 1*C*), yet differences are more apparent at lower expression levels. We believe this may be an explanation why others have seen no difference between EGFR and T790M.

EGFR signaling activates many pathways that lead to proliferative advantages. The Ras/MAPK pathway is activated via ligation of Grb2 to an activated EGFR molecule. In our 293T overexpression system, we observed that MAPK activity is increased by T790M versus WT EGFR. However, a significant effect of T790M on these downstream kinases was not evident in our HBEC3 system, which might suggest that other signaling/transcriptional events are responsible for the proliferative changes. Recent work has revealed transcriptional changes in cyclin D1, may be a key response to gefitinib resistance by T790M and susceptibility to the irreversible inhibitor CL-387,785 (14). We did not observe changes in cyclin D1 levels in the HBEC3 system and believe other, yet unidentified, mechanisms exist to account for the proliferative advantages caused by mutant EGFR.

A recent study, using a mutant-enriched PCR system, revealed the presence of the T790M mutation as a minor clone in non–small cell lung cancer tumors (19). It is suggested that these clones are selected for during gefitinib treatment and are enriched in the resistant tumor but do not provide the main proliferative function for oncogenesis. Our observations suggest that T790M as a germline mutation (i.e., in all cells) may provide the mild proliferative push for lung cancer development.

Many lines of evidence suggest the existence of a limited number of genetic factors that control susceptibility to lung cancer (20–22). However, due to a high-case fatality rate (5-year survival rate of 15%), obtaining biospecimen samples for DNA analysis is particularly difficult (23). A collaborative effort of GELCC has accrued DNA from families with lung, throat, and laryngeal cancers since the early 1990s. At present, 771 families with three or more first-degree relatives affected with lung cancer have been collected, of which 11% have sufficient family-wide biospecimen availability for any future studies. Subsequent sequencing of 237 families with predisposition to lung cancer, 45 BAC tumors and 60 fresh-frozen resected tumors, did not reveal any mutations, suggesting that T790M is likely a rare mutation. Nevertheless, our data provide a basis for *EGFR-T790M* as a rare susceptibility allele in human lung cancer.

Acknowledgments

Received 1/17/2007; revised 2/20/2007; accepted 3/9/2007.

Grant support: NIH grants U01CA76293 (Genetic Epidemiology of Lung Cancer Consortium), R01CA058554, R01CA093643, R01CA099147, R01CA099187, R01ES012063, R01ES013340, R03CA77118, R01CA80127, P30ES06096, P50CA70907 (Specialized Program of Research Excellence), and N01HG65404 and Department of Defense VITAL grant. This study was supported in part by NIH, the Intramural Research Programs of National Cancer Institute, and National Human Genome Research Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank W. Pao (Memorial Sloan-Kettering Cancer Center, New York, NY) for providing mammalian expression plasmids encoding for human EGFR, EGFR-T790M, and EGFR-L858R; R. Lubet (Chemoprevention Branch, National Cancer Institute, Bethesda, MD) for providing gefitinib; and J. Clark, Q. Chen, and M. Watson for their assistance in various aspects of this work.

References

1. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
2. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.
3. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 2004; 101:13306–11.
4. Kobayashi S, Boggon TJ, Dayaram T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med 2005;352:786–92.
5. Pao W, Miller VA, Politi KA, et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med 2005;2:e73.
6. Kwak EL, Sordella R, Bell DW, et al. Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. Proc Natl Acad Sci U S A 2005; 102:7665–70.
7. Deininger M, Buchdunger E, Druker BJ. The development of imatinib as a therapeutic agent for chronic myeloid leukemia. Blood 2005;105:2640–53.
8. Bell DW, Gore I, Okimoto RA, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet 2005;37:1315–6.
9. Blencke S, Ullrich A, Daub H. Mutation of threonine 766 in the epidermal growth factor receptor reveals a hotspot for resistance formation against selective tyrosine kinase inhibitors. J Biol Chem 2003;278:15435–40.
10. Blencke S, Zech B, Engkvist O, et al. Characterization of a conserved structural determinant controlling protein kinase sensitivity to selective inhibitors. Chem Biol 2004;11:691–701.
11. Sato M, Vaughan MB, Girard L, et al. Multiple oncogenic changes (K-RAS(V12), p53 knockdown, mutant EGFRs, p16 bypass, telomerase) are not sufficient to confer a full malignant phenotype on human bronchial epithelial cells. Cancer Res 2006;66:2116–28.
12. Bailey-Wilson JE, Amos CI, Pinney SM, et al. A major lung cancer susceptibility locus maps to chromosome 6q23-25. Am J Hum Genet 2004;75:460–74.
13. Rojas M, Yao S, Lin YZ. Controlling epidermal growth factor (EGF)-stimulated Ras activation in intact cells by a cell-permeable peptide mimicking phosphorylated EGF receptor. J Biol Chem 1996;271:27456–61.
14. Kobayashi S, Shimamura T, Monti S, et al. Transcriptional profiling identifies cyclin D1 as a critical downstream effector of mutant epidermal growth factor receptor signaling. Cancer Res 2006;66:11389–98.
15. Kosaka T, Yatabe Y, Endoh H, et al. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res 2004;64: 8919–23.
16. Yamamoto M, Kurosu T, Kakihana K, Mizuchi D, Miura O. The two major imatinib resistance mutations E255K and T315I enhance the activity of BCR/ABL fusion kinase. Biochem Biophys Res Commun 2004;319:1272–5.
17. Garcia de Palazzo IE, Adams GP, Sundareshan P, et al. Expression of mutated epidermal growth factor receptor by non-small cell lung carcinomas. Cancer Res 1993;53: 3217–20.
18. Moscatello DK, Holgado-Madruga M, Godwin AK, et al. Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors. Cancer Res 1995;55:5536–9.
19. Inukai M, Toyooka S, Ito S, et al. Presence of epidermal growth factor receptor gene T790M mutation as a minor clone in non-small cell lung cancer. Cancer Res 2006;66:7854–8.
20. Sellers TA, Bailey-Wilson JE, Elston RC, et al. Evidence for mendelian inheritance in the pathogenesis of lung cancer. J Natl Cancer Inst 1990;82:1272–9.
21. Yang P, Schwartz AG, McAllister AE, Swanson GM, Aston CE. Lung cancer risk in families of nonsmoking probands: heterogeneity by age at diagnosis. Genet Epidemiol 1999;17:253–73.
22. Chen PL, Sellers TA, Bailey-Wilson JE, Rothschild H, Elston RC. Segregation analysis of smoking-associated malignancies: evidence for mendelian inheritance. Am J Hum Genet 1991;49.
23. Bunn PAJ, Soriano A. New therapeutic strategies for lung cancer: biology and molecular biology come of age. Chest 2000;117:163–8S.

Cancer **Prevention** Research

Impact of Smoking Cessation on Global Gene Expression in the Bronchial Epithelium of Chronic Smokers

Li Zhang,[1] J. Jack Lee,[1] Hongli Tang,[2] You-Hong Fan,[2] Lianchun Xiao,[1] Hening Ren,[2]
Jonathan Kurie,[2] Rodolfo C. Morice,[2] Waun Ki Hong[2] and Li Mao[2]

Abstract Cigarette smoke is the major cause of lung cancer and can interact in complex ways with drugs for lung cancer prevention or therapy. Molecular genetic research promises to elucidate the biological mechanisms underlying divergent drug effects in smokers versus nonsmokers and to help in developing new approaches for controlling lung cancer. The present study compared global gene expression profiles (determined via Affymetrix microarray measurements in bronchial epithelial cells) between chronic smokers, former smokers, and never smokers. Smoking effects on global gene expression were determined from a combined analysis of three independent data sets. Differential expression between current and never smokers occurred in 591 of 13,902 measured genes ($P < 0.01$ and >2-fold change; pooled data)—a profound effect. In contrast, differential expression between current and former smokers occurred in only 145 of the measured genes ($P < 0.01$ and >2-fold change; pooled data). Nine of these 145 genes showed consistent and significant changes in each of the three data sets ($P < 0.01$ and >2-fold change), with eight being down-regulated in former smokers. Seven of the eight down-regulated genes, including CYP1B1 and three AKR genes, influence the metabolism of carcinogens and/or therapeutic/chemopreventive agents. Our data comparing former and current smokers allowed us to pinpoint the genes involved in smoking-drug interactions in lung cancer prevention and therapy. These findings have important implications for developing new targeted and dosing approaches for prevention and therapy in the lung and other sites, highlighting the importance of monitoring smoking status in patients receiving oncologic drug interventions.

Chronic cigarette smoking is the major cause of lung cancer and remains so for years even after smoking cessation (1, 2). Therefore, the development of agents for controlling lung cancer generally targets, virtually by default, current and former heavy smokers. Smoking status, however, seems to influence response to various chemopreventive and chemotherapeutic agents and clinical outcomes of their use (3, 4). Three large randomized clinical trials to prevent lung cancer—the Alpha-Tocopherol, Beta-Carotene Prevention Study (5), Carotene and Retinol Efficacy Trial (6), and Lung Intergroup Trial (7)—showed that current heavy smokers had harmful interactions (higher lung cancer mortality, incidence, and recurrence) with preventive agents (versus control arms); agent effects in former smokers were generally neutral and were not readily interpretable in never smokers because of the exclusion or very limited number of these patients in these trials. Certain lung cancer therapy regimens have been shown to be less effective in current smokers than in former and never smokers (8, 9). Smoking can stimulate the metabolic clearance of targeted anticancer therapies, undoubtedly diminishing therapeutic benefit (9, 10). These data highlight the importance of understanding the biological effect of chronic smoking on lung tissue.

To understand why smokers and former smokers have differential responses to agents for preventing or treating lung cancer, we analyzed and compared global gene expression profiles in three independent cancer-free cohorts comprising current, former, and never smokers.

Materials and Methods

Study population

This study included current smokers, former smokers, and never smokers with no evidence of cancer and collected from separate, independent studies conducted at The University of Texas M. D. Anderson Cancer Center (MDACC; two studies) and the Boston Medical Center (BMC; one study). The three data sets associated with the three studies are called MDACC-1, MDACC-2, and BMC throughout this article. Former smoking was defined as having quit smoking for at least

Authors' Affiliations: Departments of [1]Bioinformatics and Computational Biology and [2]Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Received 12/22/2007; revised 02/18/2008; accepted 02/21/2008.

Grant support: National Cancer Institute grant P01 CA91844 and Department of Defense grant W81XWH-04-1-0142.

Note: Supplementary data for this article are available at Cancer Prevention Research Online (http://cancerprevres.aacrjournals.org/).

Requests for reprints: Li Zhang, M. D. Anderson Cancer Center, Houston, TX 77030. E-mail: Lzhangli@mdanderson.org, or Li Mao, M. D. Anderson Cancer Center. E-mail: lmao@mdanderson.org.

©2008 American Association for Cancer Research.

doi:10.1158/1940-6207.CAPR-07-0017

12 months before study entry. Participants included in the MDACC-1 and MDACC-2 data sets came from the placebo arm of an ongoing chemoprevention trial done at M. D. Anderson Cancer Center. All MDACC subjects were clinically free of cancer at enrollment and underwent a bronchoscopy at baseline. Bronchial brushes were done at six predetermined sites including the entry area at each of the five main lobes and the carina, as previously described (11). The study was approved by the MDACC Institutional Review Board, and all MDACC participants gave signed informed consent. BMC data set participants had a bronchoscopy at the BMC and were analyzed in a previously reported study (12) as well as in the current study. Potential subjects for the MDACC or BMC data sets in the current study were excluded if their specimen images (produced as discussed below in "cRNA preparation and microarray hybridization") had defects, evidence of blood contamination, or other problems that did not meet image quality criteria applied consistently across all three data sets. All MDACC patients had smoking history, with average pack-years of 40.6 (±13). Their average age is 58 (±8) years. Sixty-one percent of them are male and 78% of them are White. More details about the demographic data of MDACC data sets can be found in Supplementary Table S2.

Bronchial brush processing and RNA extraction

For samples in MDACC-1 and MDACC-2, brushes were placed on bronchoscopy in 3 mL of plain DMEM culture (Life Technologies, Inc.) in sterile tissue culture tubes and stored at 4°C for processing the same day. The tubes were vortexed lightly to detach cells from the brushes. After removal of the brush from the tube, the cell suspension was centrifuged at 2,500 rpm for 5 min. Cell pellets were then washed with 2 mL of PBS twice, and an aliquot of material was saved at −80°C until RNA extraction. For the microarray analysis, cells from the six brushing sites of the same individual were pooled together for RNA extraction. We used TRIzol reagent (Invitrogen) for total RNA extraction according to the manufacturer's protocol, with a yield of 1 to 4 μg of total RNA per sample. Integrity of the RNA was confirmed by running it on an RNA 6000 Neno LabChip (Agilent Technologies). The samples in BMC data set were processed similarly as described (12), except that a single-round amplification protocol was used.

cRNA preparation and microarray hybridization

The first and second cDNA strands were synthesized as previously described (13). The first reverse transcription was done in the absence of biotin-labeled ribonucleotides, resulting in unlabeled cRNA, which was then used as starting material for the second cycle. In the second cycle, the first and second cDNA strands were synthesized. The second transcription was done in the presence of biotin-labeled-ribonucleotides, resulting in labeled cRNA. The cRNA was fragmented and checked by gel electrophoresis, as reported earlier (13). The Affymetrix GeneChip system was used for hybridization, staining, and imaging of the probe arrays. Hybridization cocktails of 300 μL, each containing 15 μg of cRNA and exogenous hybridization controls, were prepared as previously described and hybridized to U133A or U133A plus GeneChips (Affymetrix) overnight at 42°C. Hybridized fragments were detected with streptavidin linked to phycoerythrin (Molecular Probes). GeneChips were scanned and imaged using Affymetrix Microarray Analysis Suite version 5.0.

Microarray data normalization

There were two array types used in this study: U133A and U133 Plus 2. The U133A array contains ~500,000 distinct probe features interrogating 18,400 human transcripts and variants, including 13,902 well-characterized genes. The U133 plus 2.0 array contains all probe features that are on the U133A array. In addition, there are 9,921 new probe sets representing 6,500 new genes. To facilitate straightforward comparison of the data, we used only the probes that are common to both array types. We also ignored data of MM probes. We used PM probes common to U133A and U133 Plus 2 arrays to perform quantile normalization on the probe level data (14). The procedure was done so that the distributions of the probe signal intensities of a sample are identical for all samples within a data set. Then, we used PDNN model (15) to quantify the gene expression values from the normalized probe signal intensity data. We then applied median-centering normalization on the probe set level data so that the median of expression values of a sample was made to be the same for all the samples in all of the data sets.

Identification of differentially expressed genes

Differential expression was identified to be similar to that described by Wang et al. (16). We used Z values (defined below) to assess differentially expressed genes between current and never or former smokers, in whom the magnitude of the Z values is assumed to represent the effect of smoking cessation. For a given data set containing n_A and n_B samples in groups A and B, respectively, we compute the following test statistic Z for each probe set:

$$Z = D/\sigma \qquad (A)$$

where D is average difference between the log expression values between A and B groups. σ is the estimated standard deviation (SD) of D:

$$\sigma = \sqrt{\sigma_A{}^2/n_A + \sigma_B^2/n_B} \qquad (B)$$

where σ_A^2 and σ_B^2 are estimated variances of log expression values in groups A and B, respectively. These variances were estimated using Loess fit between the mean log expression values and the SD of the log expression values. The underlying assumption is that the mean and the SD are related by a smooth function, which allows the analysis method to treat the SD as if it were known.

Combining Z values from different data sets

The Z values obtained from the three data sets can be combined using the following formula:

$$Z = (Z_1/\sigma_1^2 + Z_2/\sigma_2^2 + Z_3/\sigma_3^2)/\omega^2 \qquad (C)$$

$$\omega^2 = 1/\sigma_1^2 + 1/\sigma_2^2 + 1/\sigma_3^2 \qquad (D)$$

where Z_1, Z_2, and Z_3 were calculated using Eq. A from MDACC-1, MDACC-2, and BMC data sets, respectively; σ_1, σ_2, and σ_3 were calculated using Eq. B from MDACC-1, MDACC-2, and BMC data sets, respectively.

The test statistic Z is supposed to form a T distribution if the log expression values are normally distributed. However, the observed data slightly deviate from the normal distribution because they contain more extreme values. Consequently, the significance of Z can be overestimated.

To alleviate the bias due to the assumption of normal distribution, we used permuted data to compute Z*. The expression values are randomly permuted for each probe set within each data set. The permutation was done 10 times to construct an empirical cumulative distribution function of Z^*. This distribution was assumed to be

Table 1. Sample sizes and array types of the microarray data sets

Data set	FS	CS	NS	Array type
MDACC-1	7	11	0	U133A
MDACC-2	8	15	0	U133 Plus 2
BMC	9	30	19	U133A

Abbreviations: FS, former smoker; CS, current smoker; NS, never smoker.

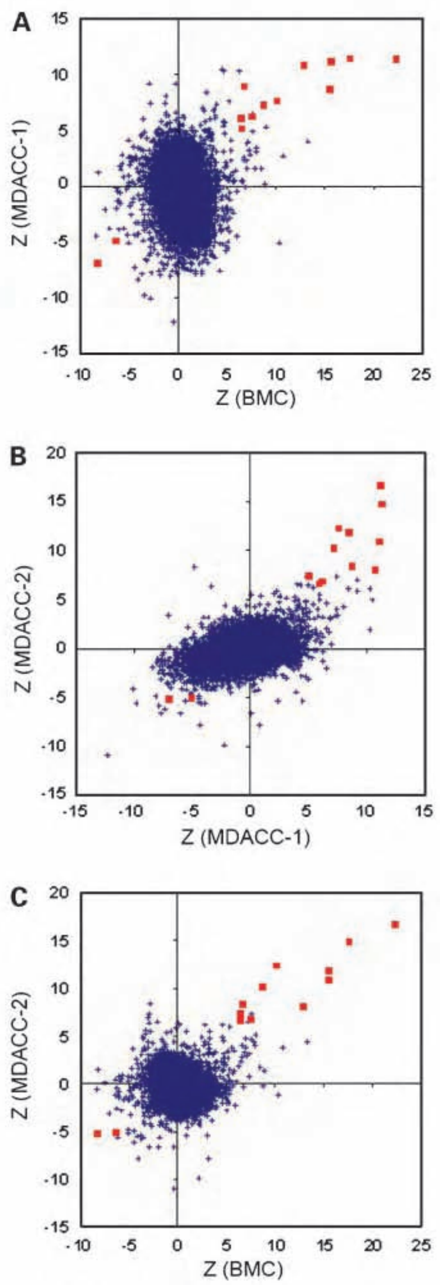


Fig. 1. Comparison of *Z* values obtained from the three data sets (BMC, MDACC-1, and MDACC-2). Each point in these scatter plots represents a probe set. The probe sets with absolute *Z* values >5 in all three data sets are shown in red. Detailed data on these probe sets are in Table 2.

the distribution of *Z* values under null hypothesis (i.e., no differential expression), and it was used to estimate the *P* values and the false discovery rate associated with *Z* values. The permutation was done within each of the data sets, but never across the data sets. Note that other than the permutation step, our method is the same as that described by Wang et al. (16).

Results

Our overall study population numbered 99 individuals, composed of 56 current smokers, 24 former smokers, and 19 never smokers from three independent data sets (Table 1). The MDACC-1 and MDACC-2 data sets included 41 chronic smokers (26 current, 15 former) enrolled in an ongoing chemoprevention trial at M. D. Anderson Cancer Center. All 41 of these subjects had at least a 20-pack-year smoking history. Demographic characteristics of the MDACC cohorts are included in Supplementary Table S1. The BMC data set was composed of 75 current, former, and never smokers. Never smokers with significant environmental cigarette exposure and subjects with respiratory symptoms or who regularly use inhaled medications were excluded. We selected 58 members of the BMC cohort for the present analysis (Table 1) and excluded 17 subjects. Exclusions from either the BMC or MDACC data sets were based on image quality criteria applied consistently across all three data sets.

First, we determined *Z* values (defined in Materials and Methods) in the three data sets separately. Then, we compared the *Z* values from each data set and, as shown in Fig. 1, we found that the genes with the most significant differential expressions (shown in red) are similar among the three data sets. The largest *Z* values mostly are located in the first and third quadrants in the scatter plots of Fig. 1, indicating that these

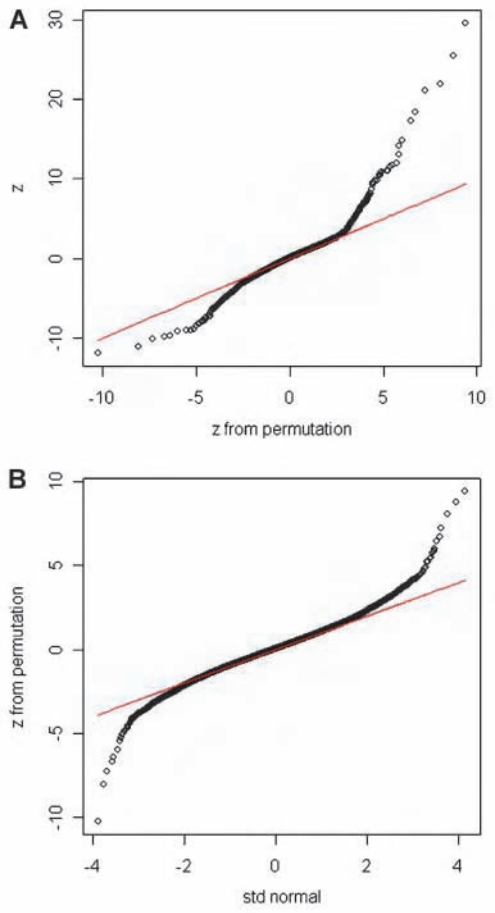


Fig. 2. Quantile-quantile plots of *Z* values. *A*, quantile of *Z* values versus quantile of *Z* values obtained from permuted data. *B*, quantile of *Z* values from permutation data versus quantile values of standard normal distribution.

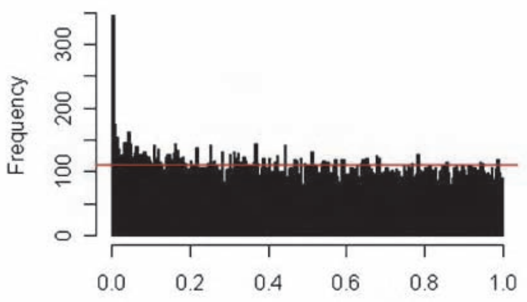


Fig. 3. Histogram of *P* values in search of differential expression between current and former smokers. Based on BUM estimate, 345 probe sets were identified as differentially expressed with a false discovery rate of 32%. One hundred seventy-six of the 345 probe sets have a fold change >2. Detailed gene information on the 176 probe sets (145 genes) is provided in Supplementary Table S2. The *P* values were evaluated on the basis of the combined *Z* values from the three data sets and the combined *Z* values from the permuted data.

changes in gene expression are consistent among the three data sets.

To further assess the statistical significance of the changes in gene expression between former and current smokers, we used quantile-quantile plots (Fig. 2) of *Z* values and BUM plots (Fig. 3) to evaluate the distribution of *P* values. Figure 2A compares the quantiles of *Z* values calculated from combining all three microarray data sets and the quantiles of Z_p (*Z* values calculated from permuted data). With permuted data, *Z* values are bounded between −10 and 10. The *Z* values from observed data contain clear outliers >10. Without differential expression, the data points in Fig. 2A should be close to the diagonal line (shown in red). Ideally, if the gene expression data obey normal distributions and are independent from each other, we would expect values of Z_p to form a standard normal distribution. However, Fig. 2B shows that Z_p's have wider ranges than that from standard normal distribution. Consequently, we used the distribution of Z_p as that from the null hypothesis (no differential expression between former and current smokers) to compute the *P* values of *Z* instead of using the standard normal distribution.

The BUM plot (17) presented a histogram of the *P* values. Under the null hypothesis, the *P* values should form a uniform distribution. The sharp spike at the left side of Fig. 3 represents the effects of differential expression contradicting the null hypothesis. The uniform part of the histogram is indicated by the red line in Fig. 3. The area above the red line contains ~1,200 probe sets, which is our estimated number of genes that are differentially expressed between the former smokers and current smokers. Only a subset of these genes is identifiable, however. According to the BUM method (17), we found 345 probe sets that were differentially expressed at a *P* value of <0.01, for which the false discovery rate was estimated to be 32%. Of the 345 probe sets, 176 have a >2-fold difference in expression (details of these 176 probe sets are shown in Supplementary Table S2). These 176 probe sets represent 145 nonredundant significantly differentially expressed genes (>2-fold change; $P < 0.01$). These 145 genes include 9 genes (Table 2) with consistent and significant changes in each of the three data sets ($P < 0.01$; >2-fold change). Eight of the nine genes are down-regulated after smoking cessation; one is up-regulated. To test the general accuracy of our microarray measurements, we compared them with reverse transcription-PCR measurements of a selected panel of genes, finding that the reverse transcription-PCR and microarray measurements were highly correlated, 96% [e.g., in the case of *ALDH3A1* (Supplementary Fig. S1), which is the gene with the largest change between former and current smokers (Table 2)]. Furthermore, although not calculated, the false discovery rate for the subset of 176 probe sets

Table 2. Genes with consistent fold changes >2 in each ($P < 0.01$) and across ($P \leq 0.0001$) the three data sets

Gene	Fold changes				*P* (comb.)	Full name	RefSeq	Probe set
	BMC	MDACC-1	MDACC-1	Comb.				
ALDH3A1	6.9	9.4	4.0	6.2	0.0000	aldehyde dehydrogenase 3 family, member A1	NM_000691	205623_at
CYP1B1	4.2	5.7	6.7	4.9	0.0000	cytochrome *P*450, member 1B1	NM_000104	202436_s_at
MUC5AC	2.2	9.6	3.0	3.5	0.0000	mucin 5AC, oligomeric mucus/gel-forming	XM_001130382	214385_s_at
AKR1C2	3.3	4.2	3.5	3.5	0.0000	aldo-keto reductase family 1, member C2	NM_001354	209699_x_at
AKR1B10	3.2	4.2	3.8	3.5	0.0000	aldo-keto reductase family 1, member B10	NM_020299	206561_s_at
AKR1C1	2.8	4.0	3.3	3.2	0.0000	aldo-keto reductase family 1, member C1	NM_001353	204151_x_at
NQO1	2.8	2.3	2.4	2.6	0.0001	NAD(P)H dehydrogenase, quinone 1	NM_000903	210519_s_at
AKR1C3	2.5	2.1	3.1	2.5	0.0000	aldo-keto reductase family 1, member C3	NM_003739	209160_at
SCGB1A1	−2.0	−2.4	−2.6	−2.4	0.0001	secretoglobin, family 1A, member 1 (uteroglobin)	NM_003357	205725_at

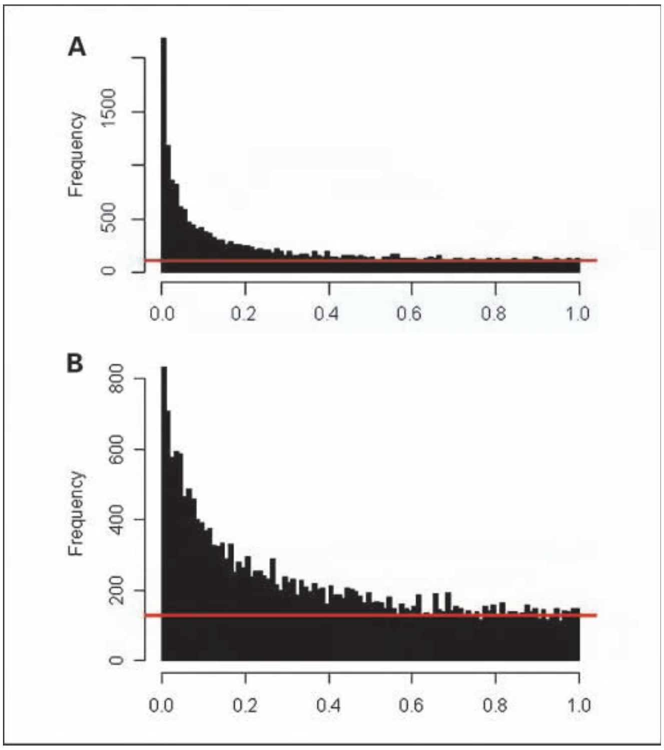


Fig. 4. Histograms of *P* values in search of differential expression between never smokers and current smokers (*A*) and between former smokers and never smokers (*B*). Only data from BMC data set were used the plots. False discovery rates were estimated to be 5% and 16% for $P < 0.01$ in *A* and *B*, respectively.

should be lower than that (32%) estimated for the 345 probe sets, and the false discovery rate for the nine changed genes that were validated across three data sets should be lower still because each subset adds new criteria that increase reliability.

For comparison, we also examined differential gene expression between current and never smokers (Fig. 4A). Similar to that in Fig. 3, the peak volume above the red line represents the number of differentially expressed genes, which is ~11,000 probe sets. This number is >9 times greater than the number detected in the comparison between former and current smokers (Fig. 3). We found 591 nonredundant genes with statistically significant changes (>2-fold change and $P < 0.01$) in pooled data of the three data sets, a group that is >4 times larger than the group of such differentially expressed genes detected in the comparison between current and former smokers. Of the 145 significantly changed genes between current and former smokers, 77 are consistent with, and 68 are not consistent with, the 591 such genes between current and never smokers (Supplementary Table S2). The nine genes with consistent and significant changes between former and current smokers in each of the three data sets are in the subset of 77 common, significantly changed genes. Figure 4B compares former smokers with never smokers.

The scope of differential expressions in Fig. 4A is much larger than that in Fig. 3, which may be due to differences in sample size. A principal component analysis (Fig. 5), however, supported the conclusion that the larger differential expression in Fig. 4A compared with that in Fig. 3 is not simply due to sample size. The gene expression profile of each patient is represented by its two main principal components. Two distinct clusters emerge in Fig. 5, and the cluster to the left (Comp1 <−10) contains mostly never smokers. The right-side cluster is predominated by a mixture of current and former smokers, which supports the conclusion that former smokers are more similar to current smokers than to never smokers.

Discussion

In probing 13,902 genes, we found that 591 were differentially expressed in current versus never smokers and that only 145 of these 591 (25%) were also differentially expressed in current versus former smokers. Among these 145 genes, 9 were significantly differentially expressed (8 overexpressed, 1 underexpressed; Table 2) by >2-fold in current versus former smokers in each ($P < 0.01$) and in the pooled data ($P < 0.0001$) of the three data sets (two MDACC, one BMC) included in this study. Therefore, our present study pinpoints and validates nine differentially expressed genes in former versus current smokers.

Seven of the eight validated genes overexpressed in current smokers—CYP1B1, four AKRs, ALDH3A1, and NQO1 (Table 2)—are involved in drug and/or carcinogen metabolism (9, 10, 18–27). Polycyclic aromatic hydrocarbons in tobacco smoke are known to bind to and activate the aryl hydrocarbon receptor and thus induce CYP1B1 (10). CYP1B1 expression is of special interest because it may contribute both to increased drug metabolism and to carcinogenesis of the aerodigestive tract (1, 18–20). The metabolic clearance of docetaxel, tamoxifen, gefitinib, erlotinib, and other cancer prevention and therapy drugs is enhanced by YP1B1 (9, 21–23). Up-regulation of CYP1B1 and the six other validated overexpressed metabolizing genes by smoking is likely involved in the adverse interactions between smoking and drugs for lung cancer prevention and therapy; smoking cessation down-regulates these gene expressions and thus may reduce or eliminate the adverse drug interactions.

Four of the eight most-up-regulated genes we detected in current smokers (Table 2) are members of the AKR

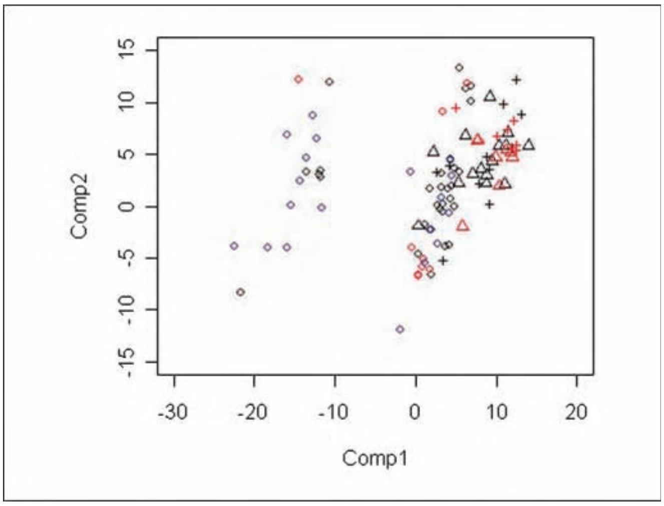


Fig. 5. Principal component analysis. The two main principal components were used to visualize the relationships among patients with different smoking status. Each point represents a patient. Current smokers were shown in black, former smokers in red, and never smokers in blue. BMC data were shown in circles, MDACC-1 data in pluses, and MDACC-2 data in triangles.

family (AKR1C1, AKR1C2, AKR1C3, and AKR1B10; ref. 28). AKR1B10 is overexpressed in non–small-cell lung cancer and squamous metaplasia in association with smoking (24, 26). AKR1C1, AKR1C2, and AKR1C3 are known to be involved in tobacco carcinogen and/or drug metabolism. AKR1C1 overexpression is correlated with a poor prognosis of non–small-cell lung cancer and is associated with chemotherapeutic drug resistance (25, 29). Data suggest that a potential role of AKR1B10 in retinoic acid signaling (30) may be a factor in the negative effects of retinoic acid (retinoids) and its relative β-carotene in smokers in chemoprevention trials (3–7). Several studies have shown that overexpression of AKR1C1, AKR1C2, or AKR1C3 contributes to the resistance of various tumor types, including lung cancer, to cisplatin-based chemotherapy (25, 31–33).

Various biases can produce inconsistencies between similar data sets. These biases can stem from differences in age, race, sex, smoking history, and sample processing. Regarding sample processing, for example, MDACC-1 and MDACC-2 samples involved two rounds of RNA amplification versus a single round in the BMC set. Two rounds of amplification are known to cause loss of signals for probes that target far away from the 3′ end of mRNA sequences. The consistent changes in smoking cessation–related genes in all three independent data sets support the robustness of our present findings.

Gene expression profiling in bronchoscopy specimens offers a direct assessment of the effects of cigarette smoking in the lungs. Gene expression patterns vary greatly between individuals, however, because of genetic variations and different environmental influences. A report by Spira et al. (on a relatively broad array of differentially expressed metabolizing and antioxidant genes in current versus former smokers; ref. 12) provided us with the opportunity to increase the robustness of our gene expression analyses by adding the BMC data set to our MDACC data sets. As we prepared our present results for publication, the Spira group published another report (34) that extended their earlier study, as do the complementary and confirmatory findings we report here. We were able to identify the specific drug-metabolizing genes involved in smoking-drug interactions, including the overexpressed genes in current versus former smokers, because of the cross validation and increased statistical power provided by adding the BMC data set (12) to our MDACC-1 and MDACC-2 data sets. The combined effect of these reports is to increase the robustness of their interrelated findings and thus their appeal for hypothesis generation.

Our results also show that the scope of genetic changes following smoking cessation is much smaller than that associated with chronic smoking (Figs. 2 and 4), possibly explaining the persistent high lung cancer risk in former smokers (35). Surprisingly at the time (~10 years ago), we and others previously found in assessments limited to specific genetic alterations that smoking-related genetic changes persisted after smoking cessation in a population similar to those of MDACC-1, MDACC-2, and BMC (36, 37). Showing similar genetic alterations in current and former smokers, results of the more sophisticated global genomic profiling approach of our present and other studies are consistent with the earlier findings (12, 34).

Our findings underscore the importance of smoking status in clinical trials, showing that smoking effects on metabolizing genes potentially can interfere with drugs in standard or investigational chemoprevention or therapy not only in the lung but in other sites as well. Future research directions should include (*a*) increased monitoring of smoking status and increased smoking cessation efforts in any trial setting because of adverse smoking effects on drug uptake and metabolism, and (*b*) the development of new dosing and targeted approaches to counteract adverse smoking-drug interactions in the lung. New targeted approaches should consider the signaling pathways of drug-metabolizing genes that were validated in this study.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733–44.

2. Cancer facts and figures, 2006 [article on the Internet]. American Cancer Society 2006. Available from: http://www.cancer.org/downloads/STT/CAFF2006PWSecured.pdf].

3. Mayne ST, Lippman SM. Cigarettes: a smoking gun in cancer chemoprevention. J Natl Cancer Inst 2005;97:1319–21.

4. Gritz ER, Dresler C, Sarna L. Smoking, the missing drug interaction in clinical trials: ignoring the obvious. Cancer Epidemiol Biomarkers Prev 2005; 14:2287–93.

5. The α-Tocopherol, β Carotene Cancer Prevention Study Group. The effect of vitamin E and β carotene on the incidence of lung cancer and other cancers in male smokers. N Engl J Med 1994;330: 1029–35.

6. Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of β carotene and vitamin A on lung cancer and cardiovascular disease. N Engl J Med 1996;334:1150–5.

7. Lippman SM, Lee JJ, Karp DD, et al. Randomized phase III intergroup trial of isotretinoin to prevent second primary tumors in stage I non-small-cell lung cancer. J Natl Cancer Inst 2001;93:605–18.

8. Zhang Z, Xu F, Wang S, et al. Influence of smoking on histologic type and the efficacy of adjuvant chemotherapy in resected non-small cell lung cancer. Lung CancerEpub ahead of print.

9. Hamilton M, Wolf JL, Rusk J, et al. Effects of smoking on the pharmacokinetics of erlotinib. Clin Cancer Res 2006;12:2166–71.

10. Port JL, Yamaguchi K, Du B, et al. Tobacco smoke induces CYP1B1 in the aerodigestive tract. Carcinogenesis 2004;25:2275–81.

11. Lee JS, Lippman SM, Benner SE, et al. Randomized placebo-controlled trial of isotretinoin in chemoprevention of bronchial squamous metaplasia. J Clin Oncol 1994;12:937–45.

12. Spira A, Beane J, Shah V, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. Proc Natl Acad Sci U S A 2004;101: 10143–8.

13. Gold D, Coombes K, Medhane D, et al. A comparative analysis of data generated using two different target preparation methods for hybridization to high-density oligonucleotide microarrays. BMC Genomics 2004;5:2.

14. Bolstad BM, Irizzary RA, Astrand M, et al. A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. Bioinformatics 2003;19:185–93.

15. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. Nat Biotechnol 2003;21:818–21.

16. Wang J, Coombes KR, Highsmith WE, et al. Differences in gene expression between B-cell chronic lymphocytic leukemia and normal B cells: a meta-analysis of three microarray studies. Bioinformatics 2004;20:3166–78.

17. Pounds S, Morris SW. Estimating the occurrence of false positives and false negatives in microarray studies b approximating and partitioning the empirical distribution of *P* values. Bioinformatics 2003;19:1236–42.

18. Mahadevan B, Luch A, Atkin J, et al. Inhibition of human cytochrome *P*450 1B1 further clarifies its role in the activation of dibenzo[*a,l*]pyrene in cells in culture. J Biochem Mol Toxicol 2007;21:101–9.

19. Roos PH, Bolt HM. Cytochrome *P*450 interactions in human cancers: new aspects considering CYP1B1. Expert Opin Drug Metab Toxicol 2005;1: 187–202.

20. Purnapatre K, Khattar SK, Saini KS. Cytochrome *P*450s in the development of target-based anticancer drugs. Cancer Lett 2008;259:1–15.

21. Sissung TM, Price DK, Sparreboom A, et al. Pharmacogenetics and regulation of human cytochrome *P*450 1B1: implications in hormone-mediated tumor metabolism and a novel target for therapeutic intervention. Mol Cancer Res 2006;4:135–50.

22. Li J, Zhao M, He P, et al. Differential metabolism of gefitinib and erlotinib by human cytochrome *P*450 enzymes. Clin Cancer Res 2007;13:3731–7.

23. Rochat B, Morsman JM, Murray GI, et al. Human CYP1B1 and anticancer agent metabolism: mechanism for tumor-specific drug inactivation? JPET 2001;296:537–41.

24. Fukumoto S, Yamauchi N, Moriguchi H, et al. Overexpression of the aldo-keto reductase family protein AKR1B10 is highly correlated with smokers' non-small cell lung carcinomas. Clin Cancer Res 2005;11:1776–85.

25. Penning TM. AKR1B10: a new diagnostic marker of non-small cell lung carcinoma in smokers. Clin Cancer Res 2005;11:1687–90.

26. Woenckhaus M, Klein-Hitpass L, Grepmeier U, et al. Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers. J Pathol 2006;210:192–204.

27. Sladek NE, Kollander R, Sreerama L, et al. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens. Cancer Chemother Pharmacol 2002;49:309–21.

28. Penning TM, Drury JE. Human aldo-keto reductases: function, gene regulation, and single nucleotide polymorphisms. Arch Biochem Biophys 2007;464:241–50.

29. Hsu NY, Ho HC, Chow KC, et al. Overexpression of dihydrodiol dehydrogenase as a prognostic marker of non-small cell lung cancer. Cancer Res 2001;61:2727–31.

30. Crosas B, Hyndman D, Gallego O, Martras S, Pares X, Flynn TG. Human aldose reductase and human small intestine aldose reductase are efficient retinal reductases: consequences for retinoid metabolism. Biochem J 2003;373:973–9.

31. Deng HB, Parekh HK, Chow KC, Simpkins H. Increased expression of dihydrodiol dehydrogenase induces resistance to cisplatin in human ovarian carcinoma cells. J Biol Chem 2002;277:15035–43.

32. Deng HB, Adikari M, Parekh HK, Simpkins H. Ubiquitous induction of resistance to platinum drugs in human ovarian cervical, germ-cell and lung carcinoma tumor cells overexpressing isoforms 1 and 2 of dihydrodiol dehydrogenase. Cancer Chemother Pharmacol 2004;54:301–7.

33. Chen J, Adikari M, Pallai R, Parekh HK, Simpkins H. Dihydrodiol dehydrogenases regulate the generation of reactive oxygen species and the development of cisplatin resistance in human ovarian carcinoma cells. Cancer Chemother Pharmacol 2007. Epub ahead of print 17661040.

34. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. Genome Biol 2007;8:R201.

35. Tong L, Spitz MR, Fueger JJ, et al. Lung carcinoma in former smokers. Cancer 1996;78:1004–10.

36. Mao L, Lee JS, Kurie JM, et al. Clonal genetic alterations in the lung of current and former smokers. J Natl Cancer Inst 1997;89:857–62.

37. Wistuba II, Lam S, Behrens C, et al. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89:1366–73.

Differential roles of DR4, DR5 and c-FLIP in regulation of geranylgeranyltransferase I inhibitor-induced augmentation of tumor necrosis factor-related apoptosis-inducing ligand-induced apoptosis. Shuzhen Chen,[1] Lei Fu,[1] Shruti M. Raja,[1] Ping Yue,[1] Yuri K. Peterson,[2] Fadlo R. Khuri,[1] and Shi-Yong Sun.[1] [1]*Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia, and* [2]*Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina*

Geranylgeranyltransferase I (GGTase I) has emerged as a cancer therapeutic target. Accordingly, small molecules that inhibit GGTase I have been developed and exhibit encouraging anticancer activity both *in vitro* and *in vivo* in preclinical studies. However, the underlying anticancer mechanisms of GGTase I inhibitors remain unclear. Here we have demonstrated a novel mechanism by which GGTase I inhibition modulates apoptosis. Inhibition of GGTase I by GGTI-298 induced apoptosis and augmented tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in human lung cancer cells. GGTI-298 induced the expression of both DR4 and DR5, two cell surface death receptors for TRAIL, and downregulated the expression of c-FLIP, a key inhibitor of death receptor-induced apoptosis. Consistently, another highly selective GGTase I inhibitor, GGTI-DU40, but not its inactive analog SN-DU40, exerted similar effects. Enforced expression of c-FLIP or knockdown of DR5 expression protected cells from induction of apoptosis by the combination of GGTI-298 and TRAIL, indicating that induction of DR5 and downregulation of c-FLIP mediate augmentation of TRAIL-induced apoptosis by GGTase I inhibition. Surprisingly, blockade of DR4 induction by knocking down DR4 expression sensitized cancer cells to GGTI298/TRAIL-induced apoptosis, suggesting that DR4 induction may play an opposite role to DR5 induction in regulating GGTI298/TRAIL-induced apoptosis. The combination of GGTI-298 and TRAIL was more effective than each single agent in decreasing the levels of IκBα and p-Akt, implying that GGTI298/TRAIL activates NF-κB and inhibits Akt. Interestingly, knockdown of DR5, but not DR4, prevented GGTI298/TRAIL-induced IκBα and p-Akt reduction, suggesting that DR5 mediates reduction of IκBα and p-Akt induced by GGTI298/TRAIL. In contrast, DR4 knockdown further facilitated IκBα and p-Akt reduction by GGTI298/TRAIL, suggesting that DR4 also plays an opposite role to DR5 in regulation of GGTI/TRAIL-induced apoptotic signaling. To our knowledge, this is the first demonstration that DR4 and DR5 may play differential roles in regulation of death receptor-mediated apoptosis. Collectively, we conclude that inhibition of GGTase I with GGTI inhibitors induces DR5 and downregulates c-FLIP, leading to augmentation of TRAIL-induced apoptosis. Thus, inhibition of GGTase I can be a novel strategy for enhancing TRAIL-based cancer therapy. (Supported by the Georgia Cancer Coalition Distinguished Cancer Scholar award, Department of Defense grant W81XWH-04-1-0142-VITAL, and NIH/NCI SPORE P50 grant CA128613-01; SY Sun, and FR Khuri are Georgia Cancer Coalition Distinguished Cancer Scholars)

The eIF4E/eIF4G interaction inhibitor 4EGI-1 augments TRAIL-induced apoptosis through DR5 induction and c-FLIP downregulation independent of inhibition of cap-dependent protein translation. Songqing Fan, Yikun Li, Ping Yue, Fadlo R. Khuri, and Shi-Yong Sun. *Departments of Hematology and Medical Oncology, Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia*

Cap-dependent protein translation plays an important role in regulation of oncogenesis primarily through regulation of the expression of certain oncogenic proteins (e.g., cyclin D1 and HIF1α,). Thus inhibition of cap-dependent protein translation has emerged as an attractive therapeutic strategy. The small molecule 4EGI-1 was identified as an inhibitor of cap-dependent translation initiation by disrupting eIF4E/eIF4G association through binding to eIF4E and exhibits growth-inhibitory and apoptosis-inducing activity in cancer cells. We were interested in its therapeutic effects in human lung cancer cells. 4EGI-1 as a single agent inhibited the growth and induced apoptosis of human lung cancer cells. When combined with the death ligand tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), enhanced apoptosis-induced activity was observed. In lung cancer cells, 4EGI-1 inhibited eIF4E/eIF4G interaction, reduced the levels of cyclin D1 and HIF1α, both of which are regulated by a cap-dependent translation mechanism. Moreover, 4EGI-1 upregulated DR5 expression and downregulated c-FLIP levels. Small interfering RNA (siRNA)-mediated blockade of DR5 induction or enforced expression of c-FLIP abrogated 4EGI-1's ability to enhance TRAIL-induced apoptosis, indicating that both DR5 induction and c-FLIP downregulation contribute to enhancement of TRAIL-induced apoptosis by 4EGI-1. However, inhibition of eIF4E/eIF4G interaction by eIF4E siRNA-mediated knockdown of eIF4E effectively reduced the levels of cyclin D1 and HIF1α, but failed to induce DR5 expression, downregulate c-FLIP levels, and augment TRAIL-induced apoptosis. As well, the mTOR inhibitor rapamycin, which inhibits cap-dependent translation initiation, did not enhance TRAIL-induced apoptosis. Collectively, we conclude that 4EGI-1 augments TRAIL-induced apoptosis through induction of DR5 and downregulation of c-FLIP independent of inhibition of cap-dependent protein translation. (S. Fan. and Y. Li share first authorship; this work was supported by the Georgia Cancer Coalition Distinguished Cancer Scholar award, DOD grant W81XWH-04-1-0142-VITAL, NIH RO1 CA118450-01 and NIH SPORE P50 grant CA128613-01; SY Sun, and FR Khuri are Georgia Cancer Coalition Distinguished Cancer Scholars)

Elevated Epithelial Insulin-Like Growth Factor Expression is a Risk Factor for Lung Cancer Development

Woo-Young Kim,[1]* Quanri Jin,[1]* Seung-Hyun Oh,[1]* Edward S. Kim,[1] Youn Joo Yang,[1] Lei Feng,[2] Carmen Behrens,[1,3] Ludmila Prudkin,[1] York E. Miller,[5] J. Jack Lee,[2,4] Scott M. Lippman,[1] Waun Ki Hong,[1] Ignacio I. Wistuba,[1,3] Ho-Young Lee[1,4]

Affiliations of authors: Departments of [1]Thoracic/Head and Neck Medical Oncology, [2]Biostatistics, and [3]Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas; [4]The University of Texas Graduate School of Biomedical Sciences, Houston, Texas; [5]Department of Medicine, Denver Veterans Affairs Medical Center, University of Colorado Denver, Denver, Colorado (YEM).
*These authors contributed equally to this work.

Funding
This work was supported by National Institutes of Health grants R01 CA109520 and CA100816-01A1 (to H.-Y. Lee); and in part: by U.S. Department of Defense grant W81XWH-04-1-0142-01-VITAL (to W.K. Hong); M. D. Anderson Cancer Center Specialized Programs of Research Excellence Grant in head and neck cancer P50 CA58187 (to W.K. Hong); and a Department of Veterans Affairs Merit Review Grant (both to Y.E.M.).

Correspondence to: Ho-Young Lee, Ph.D., Department of Thoracic/Head and Neck Medical Oncology, Unit 432, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-745-0769; Fax: 713-792-0430; E-mail: hlee@mdanderson.org.

Conflict of interest: The authors declare no conflicts of interest.

Acknowledgements
We thank OSI Pharmaceuticals Inc. and Imclone respectively for providing PQIP and A12.

Running title: IGF expression and lung carcinogenesis

Abstract

Insulin-like growth factor 1 receptor (IGF-1R) signaling has been implicated in several human neoplasms. However, the role of serum levels of insulin-like growth factors (IGFs) in lung cancer risk is controversial. Our study assessed the role of tissue-derived IGFs in lung carcinogenesis. We found that IGF-1 and IGF-2 levels in bronchial tissue specimens containing high-grade dysplasia were significantly higher than in those containing normal epithelium, hyperplasia, and squamous metaplasia. Derivatives of human bronchial epithelial cell lines with activation mutation in *KRAS* (V12) or loss of p53, genetic changes frequently observed during lung carcinogenesis, overexpressed IGF-1 and IGF-2. Tobacco carcinogen (TC) 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) enhanced transformed characteristics of these cells, which were significantly suppressed by the inhibiting the action of IGF-1, IGF-2, or the insulin-like growth factor receptor (IGFR). We further determined the role of IGF expression in lung tumorigenesis using a mouse model with a lung-specific *IGF-1* transgene after exposure to TCs, including urethane or NNK plus benzo[a]pyrene (BaP). Finally, we demonstrated antitumor activities of the selective IGF-1R tyrosine kinase inhibitor *cis*-3-[3-(4-methyl-piperazin-l-yl)-cyclobutyl]-1-(2-phenyl-quinolin-7-yl) -imidazo[1,5-*a*]pyrazin-8-ylamine (PQIP) in IGF-1 transgenic mice carrying NNK/BaP-induced lung tumors. Our results demonstrate that airway epithelial cells produce IGFs in an autocrine manner, and these IGFs act jointly with TCs to stimulate lung carcinogenesis. Thus, the use of selective IGF-1R inhibitors may be a rational approach to controlling lung cancer.

The Mechanism of EGF-Induced Hyperplasia in NHTBE Cells

Jangsoon Lee, Seung-Hee Ryu, Wen-Cheng Chung, Shin-Myung Kang, and Ja Seok Koo.
Department of Thoracic/Head and Neck Medical Oncology, the University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030-4009.

Lung cancer has been reported the highest cancer mortality in both sexes worldwide. Early detection and/or development of clinically efficient novel preventive/therapeutic targets for lung cancer are urgently required to reduce the high mortality rate associated with this disease. During the development of squamous cell carcinoma (SCC) in the lung, bronchial epithelial cells exhibit a progressive series of morphologically distinct changes: hyperplasia, squamous metaplasia, dysplasia, carcinoma *in situ*, and finally invasive SCC. Here, we investigated molecular mechanisms involved in hyperplasia of bronchial epithelial cells. We demonstrated that ErbB1 ligands, including epidermal growth factor (EGF), TGF-alpha, and amphiregulin, completely disrupted apical-basal polarity and induces hyperplasia of normal human tracheobronchial epithelial (NHTBE) cells. EGF-induced hyperplasia was completely blocked by EGFR inhibitor, erlotinib, and MEK1/2 inhibitor, U0126, suggesting involvement of MEK-ERK signaling. Further studies showed that EGF substantially upregulated cyclin D1 and these inhibitors completely blocked the upregulation. Promoter analysis of cyclin D1 revealed that Ap-1 transcription factor regulates the overexpression of cyclin D1. Depletion of AP-1 component c-Jun using siRNA completely abrogated EGF-induced cyclin D1 expression and also inhibited EGF-induced hyperplasia in NHTBE cells. In conclusion, we showed that EGF induced hyperplasia of primary bronchial epithelial cells and AP-1 plays a crucial role.

Grant Support: Department of Defense VITAL grant

STAT1 Protein Frequently Overexpressed in Non-Small Cell Lung Carcinoma.

Xiaoling Li[1,2], Ximing Tang[1], Carmen Behrens[1], Wenli Dong[1], Natalie Ozburn[1], Denise M. Woods[1], Guosheng Yin[1], Waun Ki Hong[1], Cesar Moran[1], Ignacio Wistuba[1]
[1]The University of Texas MD Anderson Cancer center, Houston TX 77030, USA.
[2] Liaoning Province Tumor Hospital, Shenyang 110042, China

Despite recent advances, the prognosis of non-small-cell lung cancer (NSCLC) is still dismal. The recurrence rate of early-stage NSCLC is ~40% within five years after a potentially curative treatment. Because of the limited prognostic power of the current pathologic and clinical criteria, providing accurate molecular prediction of the clinical outcome is currently needed. Recently, using cDNA microarray and RT-PCR methods it has been reported that the expression of genes *HER3*, *LCK* (lymphocyte-specific protein tyrosine kinase), *DUSP6* (dual-specificity phosphatase 6), and *STAT1* (signal transducer and activator transduction 1) closely associate with recurrence-free and overall survival among NSCLC patients (Chen et al, New Eng J Med 2007: 356:11-20). As only HER3 protein expression has been previously reported in lung cancer, we investigated the immunohistochemical (IHC) expression of those four genes' proteins in a large set of NSCLC tissue microarray specimens (N=306; 194 adenocarcinomas and 112 squamous cell carcinomas), and correlate their expression with clinico-pathologic features, including prognosis. Protein expression was examined semi-quantitatively using both intensity and extension of staining, and a final score was calculated for each marker. HER3 and DUSP6 proteins were expressed only in the cytoplasm of tumor cells (38% and 88% of the tumors, respectively), STAT-1 expressed in both tumor and stromal cell compartments (55% of tumors in each compartment), and LCK expressed only in inflammatory stromal cells (85% of tumors). None of the markers expression correlated with patients' recurrence-free and overall survival. STAT1 expression was lost in 45% of NSCLCs and was significantly lower in patients with squamous cell carcinoma compared to adenocarcinoma (P=0.010), non-smoking history compared to smokers (P=0.001), and in more advanced TNM stages (P=0.015). Our findings point out the difficulties of validation gene expression data using protein expression analysis and that tissue-based *in situ* methodologies are important to identify the type of cells expressing specific molecular markers. We have confirmed at protein level that STAT1 is frequently lost in NSCLC tumor tissues and the pattern of immunostaining in tumor cells is compatible with tumor suppressor gene activity and may represent a novel tumor suppressor gene for this neoplasm (Supported by Grant DoD-W81XWH-04-1-0142 and W81XWH-05-2-0027).

Immunohistochemical Expression of Estrogen and Progesterone Receptors Identifies a Subset of Non-small Cell Lung Carcinomas and Correlates with *EGFR* Mutation

Maria Gabriela Raso,[1] Carmen Behrens,[2] Suyu Liu,[3] Ludmila Prudkin,[1] Natalie C. Ozburn,[2] Denise M. Woods,[1] Ximing Tang,[2] Reza J. Mehran,[4] Cesar Moran,[1,2] J. Jack Lee,[3] and Ignacio I. Wistuba[1,2]

Authors' Affiliations: Departments of [1]Pathology, [2]Thoracic/Head and Neck Medical Oncology, [3]Biostatistics, and [4]Thoracic Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

Grant information: Supported in part by grants from the Department of Defense (W81XWH-04-1-0142 and W81XWH-05-2-0027, IIW) and by the Specialized Program of Research Excellence in Lung Cancer Grant P50CA70907 (IIW) and Cancer Center Support Grant CA-16672 from the National Cancer Institute.

Running title: ER and PR expression in NSCL

Abbreviations: EGFR, epidermal growth factor receptor; NSCLC, non-small cell lung carcinoma; ER, estrogen receptor; PR, progesterone receptor.

Key words: Estrogen receptors, Progesterone receptor, *EGFR* mutation, Field cancerization

Requests for reprints:
Ignacio I. Wistuba
Department of Pathology
The University of Texas M. D. Anderson Cancer Center
1515 Holcombe Blvd., Unit 85
Houston, TX 77030-4009
Telephone: (713) 563-9184
Fax: (713) 792-0309
E-mail: iiwistuba@mdanderson.org

ABSTRACT

Purpose: To determine the frequency of estrogen receptors (ER) α and β and progesterone receptor (PR) protein immunohistochemical expression in a large set of NSCLC specimens, and to compare our results with those for some of the same antibodies that have provided inconsistent results in previously published reports.

Experimental Design: Using multiple antibodies, we investigated the immunohistochemical expression of ERα and β and PR in 317 NSCLCs placed in tissue microarrays and correlated their expression with patients' clinicopathologic characteristics, and in adenocarcinomas, with *EGFR* mutation status.

Results: ERα and β were detected in the nucleus and cytoplasm of NSCLC cells; however, the frequency of expression (nucleus: α, 5%–36%, and β, 42%–56%; cytoplasm: α, <1%–42%, and β, 20%–98%) varied among the different antibodies tested. PR was expressed in the nuclei of malignant cells in 63% of the tumors. ERα nuclear expression significantly correlated with adenocarcinoma histology, female gender, and history of never smoking ($P = 0.0048$ to <0.0001). In NSCLC, higher cytoplasmic ERα expression significantly correlated with worse RFS (HR 1.77, 95% CI, 1.12, 2.82; $P = 0.015$) in multivariate analysis. In adenocarcinomas, ERα expression correlated with *EGFR* mutation ($P = 0.0029$ to <0.0001). ERβ and PR, but not ERα, expressed in the normal epithelium adjacent to lung adenocarcinomas.

Conclusions: ERα and β and PR are frequently expressed in NSCLC. ERα expression distinguishes a subset of NSCLC that has defined clinicopathologic and genetic features. The correlation between ER and *EGFR* mutation in lung adenocarcinoma suggests that it might be important to target both pathways simultaneously in lung cancer therapy.

INTRODUCTION

Lung cancer is the most common cause of cancer mortality worldwide, with over 1 million deaths each year (1). Lung cancer includes several histological types, the most frequently occurring of which are two types of non-small cell lung carcinoma (NSCLC): adenocarcinoma and squamous cell carcinoma (2). During the last two decades, mortality rates associated with cancer have continued to decrease across all major sites in both men and women; however, the rates for lung cancer in females have continued to increase (3, 4). Despite global statistics estimating that 15% of lung cancer in men and 53% in women are not attributable to smoking (1), smoking remains the primary risk factor for lung cancer. The higher proportion of lung cancer in females who have never smoked compared with males who have never smoked suggests a possible role for gender-dependent hormones in the development of lung cancer (5).

Estrogen receptors (ER) α and β are expressed in normal lung tissue and in lung tumors in both men and women (6), yet the data are inconsistent as to whether ER expression is gender biased (6-9) or associated with NSCLC overall survival (9-11). The data reported on the immunohistochemical expression for both ER receptors in NSCLC remain controversial. ERα has been reported to be expressed in the nucleus (0%−45%) and cytoplasm (0%−73%) of malignant lung cancer cells in the cases examined (9, 10, 12, 13). The percentages for ERβ are more consistent, with 46%−60% of NSCLC cases showing only nuclear expression (9-14). Similarly, two reports suggested that progesterone receptor (PR) is frequently (47%) expressed in NSCLC tumor cells, and this expression correlated with better patient outcome (12, 15).

Several *in vitro* and *in vivo* studies have provided evidence supporting a biological role for estrogens in lung carcinogenesis by direct promotion of cell proliferation—estrogens stimulate the proliferation of NSCLC cells through estrogen receptor-mediated signaling, whereas anti-estrogens inhibit the growth of NSCLC cells (6, 7, 13, 16, 17). Estrogen can directly stimulate the transcription of estrogen-responsive genes in the nucleus of lung cells and can also transactivate growth factor-signaling pathways—the epidermal growth factor receptor (EGFR) pathway, in particular (13, 18). In estrogen stimulation of lung cancer cells, EGFR ligands are rapidly released, activating the EGFR and mitogen-activated protein kinase 1 (MAPK1) growth pathways (19). Activation of the EGFR pathway appears to play an important role in the pathogenesis and progression of NSCLC (20). In lung cancer cells, the constitutive activation of EGFR is achieved by several mechanisms, including increased production of ligands, increased levels of the receptor, and mutation of the *EGFR* tyrosine kinase domain (20-22). Of interest, EGFR protein expression is downregulated in response to estrogens and upregulated in response to anti-estrogens, suggesting that a reciprocal control mechanism exists between the EGFR and ER pathways (19).

The purpose of the current study was to determine the frequency of ERα and β and PR protein immunohistochemical expression in a large set of NSCLCsplaced in tissue microarray (TMA) specimens and to compare our results with those for some of the same antibodies that have provided inconsistent results in previously published reports (9-14). In addition, the receptor-expression results were correlated with patients' clinicopathologic features, including NSCLC histology, gender, smoking history, and patient outcome, and in adenocarcinoma with tumors' *EGFR* activating mutation status. Finally, to understand ERα and β and PR protein expression role in the early pathogenesis of lung cancer, we investigated the characteristics of ERα and β and PR protein expression in the non-malignant respiratory epithelium adjacent to tumors taken from a subset of our retrospectively reviewed lung adenocarcinoma cases.

MATERIALS AND METHODS

Case selection and TMA construction. We obtained archived, formalin-fixed, paraffin-embedded (FFPE) tissue from surgically resected (with curative intent) lung cancer specimens (lobectomies and pneumonectomies) containing tumor and adjacent normal epithelium tissues from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX), which has been approved by the institutional review board. The tissue had been collected from 1997 to 2001, and the tissue specimens were histologically examined and classified using the 2004 World Health Organization classification system (2). We selected 317 NSCLC tissue samples (201 adenocarcinomas and 116 squamous cell carcinomas) for our TMAs. TMAs were constructed using triplicate 1-mm diameter cores per tumor, and each core included central, intermediate, and peripheral tumor tissue. Detailed clinical and pathologic information, including demographics, smoking history (never- and ever-smokers), and smoking status (never, former, and current), clinical and pathologic tumor-node-metastasis (TNM) stage, overall survival (OS) duration, and time to recurrence was available for most cases (Supplementary Table 1). Patients who had smoked at least 100 cigarettes in their lifetime were defined as smokers, and smokers who quit smoking at least 12 months before their lung cancer diagnosis were defined as former smokers. Tumors were pathologic TNM stages I–IV according to the revised International System for Staging Lung Cancer (23).

To assess the immunohistochemical expression of ERα and β and PR markers in the non-malignant respiratory epithelium adjacent to lung tumors, we selected whole histology sections containing tumor and adjacent lung tissue from 64 adenocarcinomas that were included in our TMAs.

Immunohistochemical staining and evaluation. The following antibodies against ERα and β and PR were purchased: 1) ERα-1, clone 6F11, Novocastra, Leica Microsystems Inc. (Bannockburn, IL); 2) ERα-2, clone 6F11, Chemicon, Millipore Corporate (Billerica, MA); 3) ERα-3, clone HC20, Santa Cruz Biotechnology Inc. (Santa Cruz, CA); 4) ERα-4, clone 1D5, Lab Vision Corporation (Fremont, CA); 5) ERβ-1, clone H150, Santa Cruz Biotechnology (Santa Cruz, CA); 2) ERβ-2, clone 14C8, GeneTex Inc (San Antonio, TX); and 6) PR, clone SP2, Lab Vision Corporation (Fremont, CA). Details on immunohistochemistry conditions and characteristics of the antibodies are listed in Supplementary Table 2. Immunohistochemical staining was performed as follows: 5-μM FFPE tissue sections were deparaffinized, hydrated, heated in a steamer for 10 minutes with 10 mM sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% fetal bovine serum in tris-buffered saline-t for 30 min. The slides were incubated with primary antibody at an ambient temperature for 60 min for all antibodies; the exception was ERβ 14C8 (ERβ-2), which was incubated overnight at 4°C, washed with phosphate-buffered saline, and incubated with biotin-labeled secondary antibody (Envision Dual Link +, DAKO, Carpinteria, CA) for 30 min. Staining for the slides was developed with 0.05% 3', 3-diaminobenzidine tetrahydrochloride, which had been freshly prepared in 0.05 mol/L Tris buffer at pH 7.6 containing 0.024% H_2O_2, and then the slides were counterstained with hematoxylin, dehydrated, and mounted. FFPE normal breast tissue was used as the positive control. For the negative control, we used the same specimens used for the positive controls but replaced the primary antibody with phosphate-buffered saline.

Two observers (M.G.R. and I.I.W.) jointly quantified the immunohistochemical expression of ERs and PR using light microscopy (magnification 20×). Both nuclear and cytoplasmic expressions were quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0% to 100%) of reactivity. We defined the intensity categories as follows: 0 = no appreciable staining; 1+ = barely detectable staining in epithelial cells compared with the stromal cells; 2+ = readily appreciable staining; and, 3+ = dark brown staining of cells. Next, an expression score was obtained by multiplying the intensity and reactivity extension values (range, 0–300).

EGFR mutation analysis. Exons 18–21 of EGFR were polymerase chain reaction (PCR)-amplified using intron-based primers as previously described (24, 25). Approximately 200 microdissected FFPE cells were used for each PCR amplification. All PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and DNA extraction, and the variants were sequenced in both directions, as previously reported (24, 25).

Statistical analysis. The immunohistochemical expression and clinicopathologic data data were summarized using standard descriptive statistics and frequency tabulations. BLiP plots were generated to summarize the distribution of ER and PR expressions. Associations between the marker expression and patients' clinical and demographical variables (including age, sex, smoking history, histology type, and pathologic stage) were assessed using appropriate methods including the chi-square or Fisher exact test for categorical variables, and Wilcoxon rank sum or Kruskal-Wallis test for continuous variables. The Spearman rank correlation coefficient was used to estimate the correlation between immunohistochemistry markers. Kaplan-Meier survival curves for patient OS and RFS were also generated. The log-rank test was used to identify the difference between the patient groups for both overall and RFS. For univariate and multivariate analyses for immunohistochemical expressions, the Cox proportional hazard model was used. Two-sided *P* values less than 0.05 were considered statistically significant.

RESULTS

Correlation of expression of ER antibodies. We examined four commercially available antibodies against ERα: two using the same clone (6F11) and two antibodies against ERβ (Supplementary Table 2). Using the scores of expression generated from all NSCLCs we analyzed the correlation of the expression in the malignant cells for the four ERα and the two ERβ antibodies tested. All four of the ERα antibodies showed nuclear staining, and two of the four antibodies also detected expression in the cytoplasm of malignant cells (ERα-3, clone HC20, and ERα–4, clone 1D5). The two ERα clone 6F11 antibodies (ERα-1 and ERα-2), obtained from two different companies, demonstrated only nuclear staining. All four of the ERα antibodies significantly correlated with each other at nuclear expression (Spearman rank correlation: $r = 0.32$ to 0.48; $P < 0.0001$; Supplementary Table 3). Also, significant correlation was detected in the staining of the two ERα antibodies, showing cytoplasmic expression ($r = 0.43$; $P < 0.0001$). There was no statistically significant correlation between both of the ERβ antibodies examined in their nuclear expression; although they significantly correlated at their cytoplasmic expression, the correlation coefficient was very low ($r = 0.17$; $P = 0.005$).

Frequency of ER and PR expression in NSCLC specimens by histology. We analyzed the frequency of any ER and PR immunohistochemical expression (positive cases, score >0) for each antibody tested by NSCLC tumor histology, and the data are summarized in Table 1. Representative microphotographs of the expression of ER and PR with some of the antibodies tested are shown in Fig. 1. ERs and PR were detected in the nucleus of malignant cells by all of the corresponding antibodies tested. However, when expressed, the percentage of malignant cells showing staining was low in general, with an average percentage of positive expression of: 19% (range 2-90%) for ERα-1 nuclear; 13% (range 2-93%) for ERα-2 nuclear; 21% (range 1-60%)

and 19% (range 3-73%) for ERα-3 nuclear and cytoplasmic, respectively; and, 11% (range 3-97%) and 7% (range 3-30%) for ERα-4 nuclear and cytoplasmic, respectively. The average percentages of positive cells expressing ERβ were 37 % (range 3-90%) and 37% (range 3-97%) for ERβ-1 nuclear and cytoplasmic, respectively; and, 13 % (range 1-77) and 24 % (range 3-67) for ERβ–2 nuclear and cytoplasmic, respectively.

Although there are important variations in the frequency of expression between the nuclear ERα antibodies tested, adenocarcinoma histology showed significantly higher frequency of expression than squamous cell carcinomas for all ERα antibodies (*P* <0.0001–0048; Table 1). For nuclear expression of ERβ, the data obtained with both antibodies tested were relatively consistent, and the adenocarcinoma histology demonstrated a significantly higher frequency of expression than the squamous cell carcinoma did with the ERβ-2 antibody (*P* = 0.0069). Two of the ERα (ERα-3 and ERα-4) and both ERβ antibodies also detected ER expression in the cytoplasm of NSCLC cells (Table 1). While the ERβ-2 antibody was expressed in the cytoplasm of a subset of NSCLCs, the ERβ-1 antibody was expressed in nearly all of the tumors. Cytoplasmic expression—only for the ERα-3 antibody—was significantly higher in adenocarcinomas when compared with squamous cell carcinomas (*P* = 0.0064).

In the NSCLC tissues, PR expression was frequently detected in the nuclei of malignant cells only. Squamous cell carcinoma histology showed a marginally significant higher frequency of expression than that of the adenocarcinomas (*P* = 0.05; Table 1).

Correlation between ER and PR expression in NSCLC and patients' clinicopathologic features. We correlated expression of ERs and PR for each antibody tested with the patients' clinicopathologic characteristics, including histology, gender, tobacco history, and TNM pathological stage using the expression score as a continuous variable. Using this type of analysis, adenocarcinoma histology also showed a statistically significant higher nuclear expression for all ERα antibodies and for the ERβ-2 antibody than squamous histology (Table 2). Of great interest was the fact that the NSCLC tissues obtained from females and never smokers demonstrated statistically significant higher expression of nuclear ERα and β for several of the antibodies used (Table 2). No correlations between the expression of PR and the clinicopathologic characteristics were found.

We performed OS and recurrence-free survival (RFS) analyses to determine the expression of ERs and PR for each antibody tested by using specimens from 317 patients with NSCLC with a median follow-up of 6.1 years for OS and 4.2 years for RFS. No association was detected between the expression of ER and PR and OS. Of interest, any expression of cytoplasmic ERα, using ERα-4 antibody, and nuclear ERβ, using the ERβ-1 antibody, conferred to patients a significantly worse RFS in the both univariate and multivariate analysis (Fig. 2 and Table 3). However, only the cytoplasmic expression of ERα-4 correlated with worse RFS when dichotomized score being used (HR 1.77; 95% CI, 1.11–2.81; $P = 0.0156$; Table 3).

Correlation between ER and PR expression in NSCLC and tumor EGFR mutation status. Among 182 adenocarcinoma cases, *EGFR* mutations of the tyrosine kinase domain (exons 18–21) were detected in 31 (17%) cases. Most (88%) *EGFR* mutations were detected in the exons 19 and 21, and we did not find correlation between the location of the mutation and ER α and β expression. We correlated the ER and PR scores and any expression (positive cases, score >0) with *EGFR* mutation status. Interestingly, *EGFR* mutant adenocarcinomas demonstrated statistically significant higher expression than wild-type tumors of nuclear ERα, cytoplasmic ERα, and nuclear ERβ when tested with antibodies ERα-3, ERα-4, and ERβ-1, respectively (Table 4 and Fig. 3). Because there was a higher incidence of *EGFR* mutation in lung adenocarcinoma cases from patients with a history of never smoking, Asian ethnicity, or female characteristics (data not shown), we adjusted the effects of age, gender, smoking history, ethnicity, and pathological stage in the correlation of ERα and β with *EGFR* mutation status. After linear regression analysis, all the significant correlations remained statistical significant. There was no correlation between PR expression and *EGFR* mutation status.

ER and PR immunohistochemical expression in the lung respiratory airway adjacent to adenocarcinoma cases. To characterize the pattern of expression of ER and PR in the respiratory airway field in patients with lung cancer, we selected 64 adenocarcinoma cases (35 females and 29 males; 19 never smoked, 13 current smokers, and 32 former smokers), and we studied the immunohistochemical expression of ERα and β and PR in the respiratory cells lining the small bronchi (n = 35 cases), bronchioles (n = 83 cases), and alveoli exhibiting Type II cells hyperplastic changes (n = 15 cases) using the same semiquantitative scoring system used in the TMAs. For ERs, we tested the ERα-4 and ERβ-1 antibodies. From each case, we used immunohistochemistry to examine whole tissue sections from a mean of three different paraffin blocks (range 3–6) containing tumor and adjacent normal lung tissue. We found that ERα was not expressed in the airway epithelium adjacent to lung adenocarcinomas, including epithelial samples from 21 positive tumors (Supplementary Table 4). In contrast, ERβ was widely expressed in the cytoplasm of respiratory cells: 91% of bronchi, 84% of bronchioles, and 29% of the hyperplastic alveoli. ERβ nuclear immunostaining was found less frequently: 5% of bronchi, 10% of bronchioles, and none of the hyperplastic alveolar cells. Noticeably, in the bronchial cells, we identified two patterns of cytoplasmic immunostaining: a homogeneous staining in all types of bronchial cells and heterogeneous staining comprising only ciliated cells with mainly supranuclear or apical expression (Fig. 1). PR was found in the nucleus of 56% of bronchi, 61%

of bronchioles, and 33% of hyperplastic alveoli. Of interest, there was a high level of correlation (28/33 comparisons, 85%) between the expression of PR in the normal epithelium and the corresponding tumors. Twenty (95%) out of 21 cases with PR positive in the normal epithelium were detected in patients with tumors that also expressed this receptor.

DISCUSION

ERα and β frequently expressed in our NSCLC cases, and ERα expression distinguished a subset of NSCLC that has defined clinicopathologic and genetic features. Although the immunohistochemical expression of ERα and β has been reported in tumor tissue specimens from surgically resected NSCLCs, the data on the fraction of tumors expressing ER are still controversial. Previous studies on ERα immunohistochemical expression in formalin-fixed and paraffin-embedded NSCLC specimens using seven different antibodies identified nuclear expression in malignant cells in frequencies that ranged from none (10, 14) to 18% (26) and 38% (12). Similarly, in other studies, the frequency of ERα cytoplasmic expression in NSCLC ranged from 0%–3% (12, 26) to 35% (11) and 73% (27). In the current study, using four different commercially available ERα antibodies, we also identified a wide range of percentages in the frequency of NSCLCs exhibiting any expression of ERα in the nucleus (7%–54%) and in the cytoplasm (0%–42%) of tumor cells. However, in our study, when the scores of immunohistochemical expression were analyzed as continuous variables, all of the ERα antibodies significantly correlated with each other at nuclear and cytoplasmic locations.

A similar situation is observed when the ERβ immunohistochemical expression data are examined in NSCLC. Several previous studies, using six different antibodies, have reported frequencies of ERβ expression in tumors with a wide range of percentages at the nuclear location—0% (9), 34%47% (10, 12, 14), and 61%–84% (9, 11) —but not in the cytoplasm of malignant cells, where most of the studies have shown no reactivity (9, 10, 12, 14); some expression was seen in a small number of cases (6) or low frequency of expression in a large number of cases (10%) (11). In the present study using two antibodies, any ERβ nuclear expression was detected in about half (56% and 42%) of the NSCLCs, and cytoplasmic expression was found in a wider range (20%–98%) of our cases. We do not have a definitive explanation to the high levels of expression of ERβ in NSCLC cells in our study, and the discordance with previous reports. However, immunohistochemical analysis has shown the distribution of ERβ to be much more widespread than ERα (28-30). Several studies have reported that ERβ immunohistochemical expression is frequently detected in the nucleus and cytoplasm of normal respiratory cells (28). While expression has been questioned by suggestions that this observation is based on non-specific binding produced by unpurified antibodies (31), multiple reports have shown the presence of a non-nuclear pool of ERs in normal and malignant cells (32-35). Yang et al (35) used one of the same ERβ antibodies that we used (ERβ-1) and demonstrated mitochondrial localization of this receptor in several normal human and murine cells, suggesting a role for ERβ receptor in the cytoplasm of cells. Our finding of high frequency of ERβ expression, using ERb-1 antibody, in the cytoplasm of normal respiratory cells from our lung adenocarcinoma patients are consistent with these findings.

Several discrepancies were observed when we compared our results with those published previously (6, 9, 10, 12) using the same antibodies, especially for ERα. For example, our ERα-3 antibody, raised against the COOH-terminus region of the protein, detected any nuclear and cytoplasmic expressions in 54% and 42% of our NSCLC cases, respectively. Using this antibody, nuclear expression was reported in a small number of NSCLC tumors by Stabile et al (6) and in none of the 130 tumors examined by Kawai et al (10). At the cytoplasmic location of malignant cells, both studies reported positive immunostaining (6, 10), with up to 73% of cases in the study performed by Kawai et al (10).

Why these inconsistent results on the immunohistochemical expression of ER α and β occur raises a very important question. Clearly the reasons for the inconsistent results include the use of different antibodies manufactured from different clones and by different companies. Indeed, some of these antibodies have been made against different parts of the protein: full length, NH_2-terminues, and COOH-terminus regions. It has been suggested that several mRNA splicing variants of ERα have been detected in lung cancer cell lines, and antibodies raised against epitopes in the deleted exons of ER may give conflicting results (6). In addition, it is important to note that there are multiple criteria reported to assess ERα and β positivity in NSCLC tissues. Although most studies considered different levels of intensity (usually a scale 0–3+) of expression at nuclear and cytoplasmic locations combined with the percentage of malignant cells expressing a given intensity, the cutoff levels of expression vary significantly between studies (e.g., 1+ in >10% of cells; 1+ in 1%–25% of cells; >50% of cells; score “0–8”, etc) (6, 9-12, 14, 26).

Because there were different levels of ERα and β immunohistochemical expression detected using different antibodies in ours and the previous studies (6, 9-12, 14, 26), we correlated the expression of ER using all of the antibodies we tested with the patients’ clinicopathological features and the tumors’ *EGFR* mutation status. The evaluation of multiple antibodies for ER expression adds strength to our findings. In our study, we analyzed the immunohistochemical scores as continuous and dichotomized variables, and a significantly higher expression of nuclear ERα was detected with all four antibodies tested in adenocarcinoma than squamous cell carcinoma histology—three out of the four antibodies tested in tumors obtained from females compared with males and from people who had never smoked compared with smokers. The two previous studies reporting ERα nuclear expression in NSCLC, which examined a relatively large series of cases, did not address differences of expression based on histology types or patients’ clinicopathologic features (12, 26). In the NSCLC tissues that we reviewed, higher expression of ERβ correlated significantly with tumor adenocarcinoma histology and the patients' female gender for ERβ-1 antibody, and correlated with the patients' history of never smoking with the ERβ-2 antibody.

Few studies have shown inconsistent results on whether ER expression is biased to any gender using different types of specimens and assays (6-9). Schwartz et al (9), using a different antibody than ours, reported that NSCLCs obtained from females were 46% less likely to have ERβ-positive tumors than males in a multivariate analysis. In addition, mRNA expression of ERα has been reported to be significantly higher in lung tumors from women than from men (8). In a small number of NSCLC tumor tissue specimens, ERα and β gene transcripts have been found to be expressed in similar levels when comparing samples obtained from females and males (7). Adenocarcinoma of the lung, which shows a weaker association with tobacco smoking than with other types of lung cancer, is also found predominantly in women, suggesting a possible role for female hormones in the pathogenesis of this type of lung cancer (5).

In previous studies, ERβ expression in NSCLC tumors has been associated with improved survival (9-11), whereas the immunohistochemical expression of ERα has been shown to be a poor prognostic factor (9). Thus, both ERs have been proposed to play opposite roles in cell proliferation, with ERα promoting proliferation and ERβ having an anti-proliferative effect (36, 37). In our study, we did not find a correlation between OS and RFS and ERβ expression, but we did find that only the expression of cytoplasmic ERα (using one antibody) conferred to patients a significantly worse RFS, but not OS, in multivariate analysis.

Several studies have shown that estrogen signaling plays a role in the development of the epithelium in the lung and that estrogen could potentially promote lung cancer (6, 7, 13, 16, 17). Additionally, anti-estrogen drugs have been suggested to have a role in the therapy of lung cancer (6, 19). NSCLC cell lines and *in vivo* tumor xenografts have been shown to respond to estrogens, and tumor growth can be inhibited up to 40% by the anti-estrogen fulvestrant (6). In the past few years, significant advances have been made in the development of new molecularly targeted agents for lung cancer (38). The identification of the subset of patients with NSCLC who will benefit with targeted therapy is a key element in the development of personalized treatment approaches in this disease. A pilot study of combined therapy using fulvestrant and gefitinib in advanced NSCLC has shown to be well tolerated and has demonstrated some tumor responses (39). Our study results strongly suggest that NSCLC tumors obtained from patients with adenocarcinoma histology, female gender, and history of never smoking have a higher chance of expressing ERs and have the potential to respond positively to anti-estrogen therapy.

To the best of our knowledge, our study is the first to report an association between *EGFR* mutation and ERα and β expression in lung adenocarcinomas. Importantly, we have demonstrated that the correlation between ER expression and *EGFR* mutation is independent of the clinicopathological features associated with both abnormalities, such as adenocarcinoma histology, female gender, and history of never smoking (40). Based on the interactions between ER and EGFR-signaling pathways, there is evidence showing that targeting both pathways by using anti-estrogens (fulvestrant) and EGFR tyrosine kinase inhibitors (gefitinib), the antitumor effect in *in vitro* and *in vivo* lung cancer models of the drug combination is higher than in treatment with each drug alone (19). Thus, our findings of an association between the activation of both pathways further strengthens the concept of combined anti-estrogen and EGFR inhibitor therapy for a selected group of patients with lung adenocarcinoma.

Although PR expression has been reported to be present in NSCLC cell lines and tumor specimens, the data are controversial like those for ERs (11, 12, 15, 41, 42). Out of four studies reporting on immunohistochemical expression of PR in surgical resected and formalin-fixed NSCLC tissue specimens using different antibodies, there were two studies that reported a relatively high frequency of PR expression in tumors (39% and 47%) (12, 15); the remaining two reports showed no expression (11, 42). In the present study, PR was frequently (63%) detected in the nuclei of malignant NSCLC, with a trend to higher expression in squamous cell carcinoma histology. We did not find a correlation between PR and any of the clinicopathologic characteristics we studied, including survival. In contrast, Ishibashi et al reported that PR immunohistochemical expression was higher in NSCLCs obtained from females and correlated with better OS in stages I–III tumors (12). In breast cancer, transcription of the PR gene is well known to be regulated by estrogenic actions through estrogen receptors, and a positive PR status is generally regarded as one of the markers of functional estrogenic pathways. In our study, we found no statistical correlation between PR and any of the ER antibodies studied. *In vitro* and *in vivo* studies have shown that administration of progesterone inhibits the growth of PR-positive NSCLC cell lines, which is similar to what has been shown to happen in breast and endometrial carcinomas (12).

Lung cancer is believed to develop from a series of preneoplastic lesions in the respiratory mucosa, and these abnormalities are frequently extensive and multifocal throughout the respiratory epithelium, indicating a field-effect or field-cancerization phenomenon (43). Our findings of relatively frequent expression of nuclear PR and lack of expression of ERα in the normal epithelium adjacent to adenocarcinomas expressing these receptors suggest that PR, but not ERα expression, may represent a field-effect phenomenon. Of interest, all but one case with normal epithelium expression of PR showed expression of this receptor in the corresponding tumor. The frequent finding of cytoplasmic ERβ in normal epithelium may represent a constitutive expression in normal respiratory cells and is probably not related to the carcinogenesis process (35).

In summary, our findings show that ERα and β and PR are frequently expressed in NSCLC, and ER expression distinguishes a subset of NSCLC that has defined clinicopathologic and genetic features. The correlation between ER and EGFR mutation in lung adenocarcinoma suggests that it might be important to target both pathways simultaneously for lung cancer chemoprevention and therapy.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005; 55:74-108.

2. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart. Lyon: International Agency for Research on Cancer (IARC); 2004. p. 9-124.

3. Brennan P, Bray I. Recent trends and future directions for lung cancer mortality in Europe. Br J Cancer 2002; 87:43-8.

4. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008; 58:71-96.

5. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007; 7:778-90.

6. Stabile LP, Davis AL, Gubish CT, et al. Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor alpha and beta and show biological responses to estrogen. Cancer Res 2002; 62:2141-50.

7. Mollerup S, Jorgensen K, Berge G, Haugen A. Expression of estrogen receptors alpha and beta in human lung tissue and cell lines. Lung Cancer 2002; 37:153-9.

8. Fasco MJ, Hurteau GJ, Spivack SD. Gender-dependent expression of alpha and beta estrogen receptors in human nontumor and tumor lung tissue. Mol Cell Endocrinol 2002; 188:125-40.

9. Schwartz AG, Prysak GM, Murphy V, et al. Nuclear estrogen receptor beta in lung cancer: expression and survival differences by sex. Clin Cancer Res 2005; 11:7280-7.

10. Kawai H, Ishii A, Washiya K, et al. Estrogen receptor alpha and beta are prognostic factors in non-small cell lung cancer. Clin Cancer Res 2005; 11:5084-9.

11. Skov BG, Fischer BM, Pappot H. Oestrogen receptor beta over expression in males with non-small cell lung cancer is associated with better survival. Lung Cancer 2008; 59:88-94.

12. Ishibashi H, Suzuki T, Suzuki S, et al. Progesterone receptor in non-small cell lung cancer--a potent prognostic factor and possible target for endocrine therapy. Cancer Res 2005; 65:6450-8.

13. Marquez-Garban DC, Chen HW, Fishbein MC, Goodglick L, Pietras RJ. Estrogen receptor signaling pathways in human non-small cell lung cancer. Steroids 2007; 72:135-43.

14. Wu CT, Chang YL, Shih JY, Lee YC. The significance of estrogen receptor beta in 301 surgically treated non-small cell lung cancers. J Thorac Cardiovasc Surg 2005; 130:979-86.

15. Su JM, Hsu HK, Chang H, et al. Expression of estrogen and progesterone receptors in non-small-cell lung cancer: immunohistochemical study. Anticancer Res 1996; 16:3803-6.

16. Pietras RJ, Marquez DC, Chen HW, Tsai E, Weinberg O, Fishbein M. Estrogen and growth factor receptor interactions in human breast and non-small cell lung cancer cells. Steroids 2005; 70:372-81.

17. Hershberger PA, Vasquez AC, Kanterewicz B, Land S, Siegfried JM, Nichols M. Regulation of endogenous gene expression in human non-small cell lung cancer cells by estrogen receptor ligands. Cancer Res 2005; 65:1598-605.

18. Stabile LP, Siegfried JM. Estrogen receptor pathways in lung cancer. Curr Oncol Rep 2004; 6:259-67.

19. Stabile LP, Lyker JS, Gubish CT, Zhang W, Grandis JR, Siegfried JM. Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects. Cancer Res 2005; 65:1459-70.

20. Sequist LV, Lynch TJ. EGFR tyrosine kinase inhibitors in lung cancer: an evolving story. Annu Rev Med 2008; 59:429-42.

21. Herbst RS, Bunn PA, Jr. Targeting the epidermal growth factor receptor in non-small cell lung cancer. Clin Cancer Res 2003; 9:5813-24.

22. Fujimoto N, Wislez M, Zhang J, et al. High expression of ErbB family members and their ligands in lung adenocarcinomas that are sensitive to inhibition of epidermal growth factor receptor. Cancer Res 2005; 65:11478-85.

23. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997; 111:1710-7.

24. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005; 97:339-46.

25. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005; 65:7568-72.

26. Lau SK, Chu PG, Weiss LM. Immunohistochemical expression of estrogen receptor in pulmonary adenocarcinoma. Appl Immunohistochem Mol Morphol 2006; 14:83-7.

27. Kawai H, Ishii A, Washiya K, et al. Combined overexpression of EGFR and estrogen receptor alpha correlates with a poor outcome in lung cancer. Anticancer Res 2005; 25:4693-8.

28. Taylor AH, Al-Azzawi F. Immunolocalisation of oestrogen receptor beta in human tissues. J Mol Endocrinol 2000; 24:145-55.

29. Jarvinen TA, Pelto-Huikko M, Holli K, Isola J. Estrogen receptor beta is coexpressed with ERalpha and PR and associated with nodal status, grade, and proliferation rate in breast cancer. Am J Pathol 2000; 156:29-35.

30. Skliris GP, Carder PJ, Lansdown MR, Speirs V. Immunohistochemical detection of ERbeta in breast cancer: towards more detailed receptor profiling? Br J Cancer 2001; 84:1095-8.

31. Saunders PT. Does estrogen receptor beta play a significant role in human reproduction? Trends Endocrinol Metab 2005; 16:222-7.

32. Collins P, Webb C. Estrogen hits the surface. Nat Med 1999; 5:1130-1.

33. Razandi M, Oh P, Pedram A, Schnitzer J, Levin ER. ERs associate with and regulate the production of caveolin: implications for signaling and cellular actions. Mol Endocrinol 2002; 16:100-15.

34. Speirs V, Carder PJ, Lane S, Dodwell D, Lansdown MR, Hanby AM. Oestrogen receptor beta: what it means for patients with breast cancer. Lancet Oncol 2004; 5:174-81.

35. Yang SH, Liu R, Perez EJ, et al. Mitochondrial localization of estrogen receptor beta. Proc Natl Acad Sci U S A 2004; 101:4130-5.

36. Lindberg MK, Moverare S, Skrtic S, et al. Estrogen receptor (ER)-beta reduces ERalpha-regulated gene transcription, supporting a "ying yang" relationship between ERalpha and ERbeta in mice. Mol Endocrinol 2003; 17:203-8.

37. Stettner M, Kaulfuss S, Burfeind P, et al. The relevance of estrogen receptor-beta expression to the antiproliferative effects observed with histone deacetylase inhibitors and phytoestrogens in prostate cancer treatment. Mol Cancer Ther 2007; 6:2626-33.

38. Lynch TJ, Bonomi PD, Butts C, et al. Novel agents in the treatment of lung cancer: Fourth Cambridge Conference. Clin Cancer Res 2007; 13:s4583-8.

39. Traynor AM, Schiller JH, Stabile LP, et al. Pilot study of gefitinib and fulvestrant in the treatment of post-menopausal women with advanced non-small cell lung cancer. Lung Cancer 2008.

40. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006; 118:257-62.

41. Kaiser U, Hofmann J, Schilli M, et al. Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer. Int J Cancer 1996; 67:357-64.

42. Di Nunno L, Larsson LG, Rinehart JJ, Beissner RS. Estrogen and progesterone receptors in non-small cell lung cancer in 248 consecutive patients who underwent surgical resection. Arch Pathol Lab Med 2000; 124:1467-70.

43. Wistuba, II. Genetics of preneoplasia: lessons from lung cancer. Curr Mol Med 2007; 7:3-14.

Figure Legends

Fig. 1. Fig. 1. Microphotographs showing representative examples of immunohistochemical expression of estrogen receptors (ER) α (panels *A*, *B*, and *C*) and β (panels *D*, *E* and *F*) and progesterone receptor (PR; panels *G*, *H* and *I*), in tissue specimens of non-small cell lung carcinoma (NSCLC) tumors and bronchial epithelium (panels *A*, *D*, and *G*) adjacent to adenocarcinomas. The two NSCLC histologies are represented: adenocarcinoma (panels *B*, *E* and *H*), and squamous cell carcinoma (panels *C*, *F* and *I*). ER α and β expressions are shown using antibodies ER-α4 and ER-β1, respectively. Examples of nuclear expression and cytoplasmic expressions are indicated by red and blue arrows, respectively. Normal epithelia show nuclear expression of ER-β and PR, and cytoplasmic expression of ER-β. Adenocarcinomas show nuclear expression for all three markers and cytoplasmic staining for ER-β. Squamous cell carcinomas show nuclear staining for PR, and both nuclear and cytoplasmic for ER-β. The magnification of the microphotographs is 200x for normal epithelium and 400x for tumors.

Fig. 2. Kaplan-Meier curves showing recurrence free survival (RFS) of NSCLC patients for ERα cytoplasmic (panel *A*) and ERβ nuclear (panel *B*) expression.

Fig. 3. Representative examples of ERα and β immunohistochemical expression (*upper figures*) and epidermal growth factor receptor *(EGFR)* mutations (*lower figures*) in lung adenocarcinomas. Panel *A*: ERα (antibody ERα-4) positive in the nucleus of malignant cells and sequencing chromatograms showing the presence of mutant form of *EGFR* (15 bp deletion in exon 19; arrow indicates in-frame deletion mutation sequence). Panel *B*: ERβ (antibody ERβ-1)

positive in the nucleus of malignant cells and sequencing chromatograms showing the presence of mutant *EGFR* (L858R point mutation in exon 21; arrow indicates CTG to CGG mutation). Panel *C*: ER α and β (same antibodies than panels *A* and *B*) with negative expression in the malignant cells and sequencing chromatograms showing the presence of wild-type form for *EGFR* exon 19 (line indicates sequence 746 to 750).

Table 1. Frequency of ER and PR immunohistochemical expression[*] in NSCLC tissue specimens

Marker	**Location**	**Adenocarcinoma**		**Squamous cell carcinoma**		
		Number of cases	**Positive N (%)**	**Number of cases**	**Positive N (%)**	***P* value**
ERα-1	Nucleus	187	20 (11)	109	2 (2)	0.0048
	Cytoplasm	187	0	108	0	--[†]
ERα-2	Nucleus	186	84 (45)	110	23 (21)	<0.0001
	Cytoplasm	185	1 (<1)	111	0	1.000
ERα-3	Nucleus	191	16 (8)	114	0	0.0007
	Cytoplasm	190	92 (48)	114	37 (33)	0.0064
ERα-4	Nucleus	185	74 (40)	109	25 (23)	0.0028
	Cytoplasm	185	35 (19)	109	18 (17)	0.6043
ERβ-1	Nucleus	189	102 (54)	112	66 (59)	0.4022
	Cytoplasm	189	185 (98)	112	110 (98)	1.0000
ERβ-2	Nucleus	174	83 (48)	100	31 (31)	0.0069
	Cytoplasm	172	37 (22)	100	16 (16)	0.2685
PR	Nucleus	177	103 (58)	112	78 (70)	0.05
	Cytoplasm	176	0	112	0	--[†]

[*] Any expression score >0 is considered positive.

[†] Not tested.

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; NSCLC, non-small cell lung cancer; ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Table 2. Significant correlations between immunohistochemical expression of ER and PR* and NSCLC patients' clinicopathological features.

Estrogen Receptor	**Histology** **ADCA (n = 201)** **>** **SCC (n = 116)**	**Gender** **Female (n = 167)** **>** **Male (n = 150)**	**Tobacco History** **Never (n = 54)** **>** **Ever (n = 262)**
ERα-1 nucleus	0.0048	0.0051	ns[†]
ERα-2 nucleus	<0.0001	0.0109	0.0006
ERα-3 nucleus	0.0015	ns	0.0242
ERα-4 nucleus	0.0004	0.0148	0.0044
ERβ-1 nucleus	ns	ns	0.0290
ERβ-2 nucleus	0.0016	0.044	ns

*ER and PR were tested using expression score.

[†] ns = not statistically significant.

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; NSCLC, non-small cell lung cancer; ADCA, adenocarcinoma SCC, squamous cell carcinoma; ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Table 3. Multivariate RFS analysis using Cox regression model in NSCLC patients[*]

Variable	**Hazard ratio**	**95% CI of HR**		***P* value**
		Lower limit	**Upper limit**	
ER as continuous variable				
ERα-4 cytoplasm	1.05	1.01	1.08	0.0068
ERβ-1 nucleus	1.01	1.00	1.02	0.0034
Stage II vs. I	1.90	1.14	3.18	0.0145
Stage III/IV vs. I	3.17	1.98	5.08	<.0001
ER dichotomized				
ERα-4 cytoplasm: >0 vs. 0	1.77	1.11	2.81	0.0156
ERβ-1 nucleus: >0 vs. 0	1.36	0.91	2.05	0.1388
Stage II vs. I	1.79	1.08	2.99	0.0250
Stage III/IV vs. I	3.13	1.97	4.99	<.0001

*With only significant covariates.

Abbreviations: NSCLC, non-small cell lung cancer; CI, confidence interval; HR, hazard ratio; ER, estrogen receptor; ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Table 4. Significant correlations between immunohistochemical expression of ER and *EGFR* mutation status in adenocarcinoma

ER expression by antibody	***EGFR* mutation status** **Wild-type** **N positive/total (%)**	 **Mutant** **N positive/total (%)**	***P* value**
ERα-3 nucleus	9/146 (6)	7/28 (25)	0.0016
ERα-3 cytoplasm	68/146 (47)	21/27 (78)	0.0029
ERα-4 nucleus	50/143 (35)	18/27 (67)	0.0020
ERα-4 cytoplasm	20/143 (14)	13/27 (48)	<0.0001
ERβ-1 nucleus	70/145 (48)	22/27 (82)	0.00015

*ER and PR were tested using expression score.

Abbreviations: ER, estrogen receptor; EGFR, epidermal growth factor receptor; PR, progesterone receptor; ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Supplementary Table 1. Demographic, clinicopathologic, and *EGFR* mutation status data of the NSCLCs studied

Characteristic	**Number (%)**
Gender	
Female	167 (53)
Male	150 (47)
Smoking status*	
Former	160 (51)
Current	102 (32)
Never	54 (17)
Histology	
Adenocarcinoma	201 (63)
Squamous cell carcinoma	116 (37)
Stage	
I	202 (63)
II	62 (20)
III	46 (15)
IV	7 (2)
EGFR mutation†	
No	151 (83)
Yes	31 (17)

*Smoking history and status were not available for one patient with squamous cell carcinoma.

†*EGFR* mutation data was available in 182 cases.

Abbreviations: EGFR, epidermal growth factor receptor; NSCLC, non-small cell lung cancer.

Supplementary Table 2. Antibodies used for immunohistochemical analysis of ER and PR in NSCLC tissue specimens

Receptor	**Raised against**	**Clone**	**Company**	**Dilution**	**Code**[*]
ER-α	Full length	6F 11	Novocastra	1:400	ERα-1
	Full length	6F 11	Chemicon	1:50	ERα-2
	C-term	HC20	Santa Cruz	1:500	ERα-3
	N-term	1D5	Lab Vision	1:50	ERα-4
ER-β	aa 1 to 150	H150	Santa Cruz	1:100	ERβ-1
	aa 1 to 153	14C8	Genetex	1:100	ERβ-2
PR	Full length	SP2	Lab Vision	1:75	PR

[*] Code utilized to refer to each antibody in this study.

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; NSCLC, non-small cell lung cancer; ERα, estrogen receptor-α; ERβ, estrogen receptor-β; aa, aminoacid

Supplementary Table 3. Summary of the Spearman's correlation between ERα and ERβ antibodies examined**.**

Nucleus	ERα-2	ERα-3	ERα-4
ERα-1	*r* 0.41 – *P* < 0.0001	*r* 0.48 – *P* < 0.0001	*r* 0.34 – *P* < 0.0001
ERα-2	--	*r* 0.32 – *P* < 0.0001	*r* 0.44 – *P* < 0.0001
ERα-3	--	--	*r* 0.41 – *P* < 0.0001
	ERβ-2		
ERβ-1	*r* 0.02 – *P* = 0.79		

Cytoplasm	ERα-4
ERα-3	*r* 0.43 – *P* < 0.0001
	ERβ-2
ERβ-1	*r* 0.17 – *P* = 0.005

Abbreviations: ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Supplementary Table 4. Expression of ER and PR in the respiratory epithelium adjacent to lung adenocarcinomas

Tissue site	ERα[*]		ERβ[†]		PR
	Nuclear	**Cytoplasm**	**Nuclear**	**Cytoplasm**	**Nuclear**
	N + (%)	**N + (%)**	**N + (%)**	**N + (%)**	**N + (%)**
Tumor	21/51 (41)	6/51 (12)	19/47 (40)	41/47 (87)	27/36 (75)
Bronchi	0/35	0/35	2/37 (5)	34/37 (91)	18/32 (56)
Bronchioles	0/83	0/83	8/81 (10)	68/81 (84)	43/71 (61)
Alveoli	0/15	0/15	0/14	4/14 (29)	5/15 (33)

[*] Antibody ERα–4.

[†] Antibody ERβ–1.

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; ERα, estrogen receptor-α; ERβ, estrogen receptor-β.

Figure 1

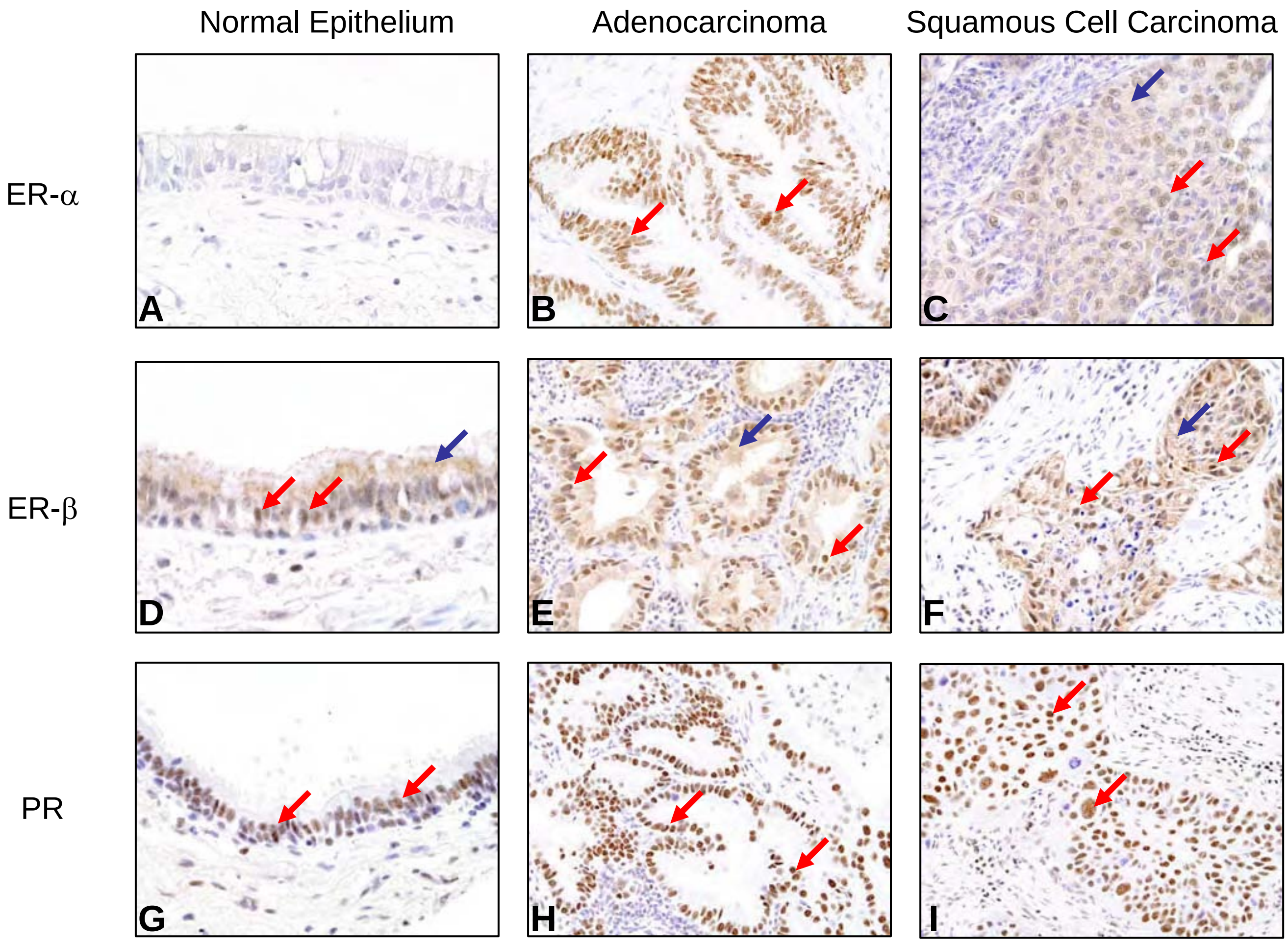
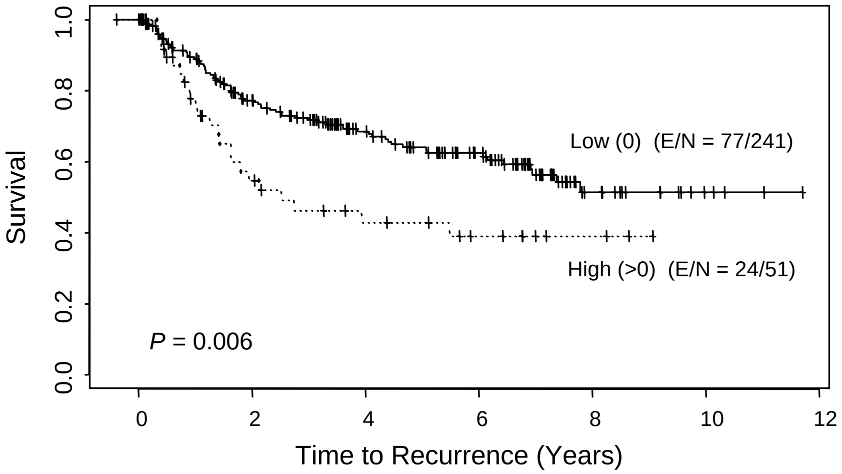

Normal Epithelium
Adenocarcinoma
Squamous Cell Carcinoma
ER-α
ER-β
PR
A
B
C
D
E
F
G
H
I

Figure 2

A. ER-α Cytoplasmic Expression



B. ER-β Nuclear Expression

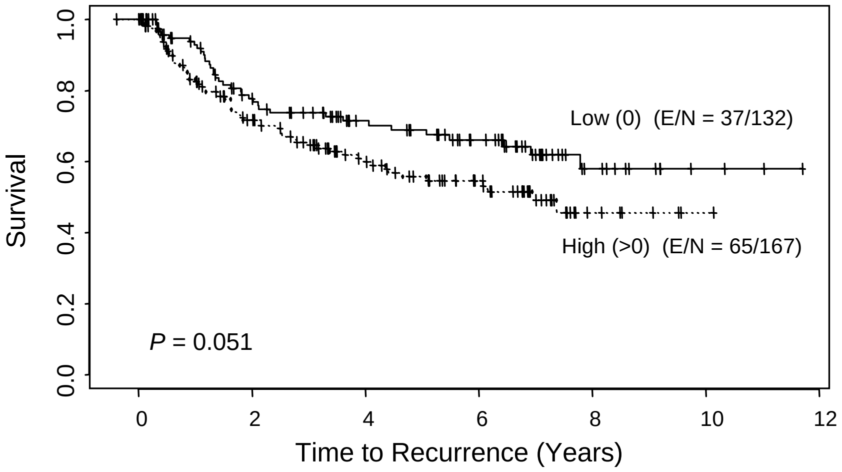
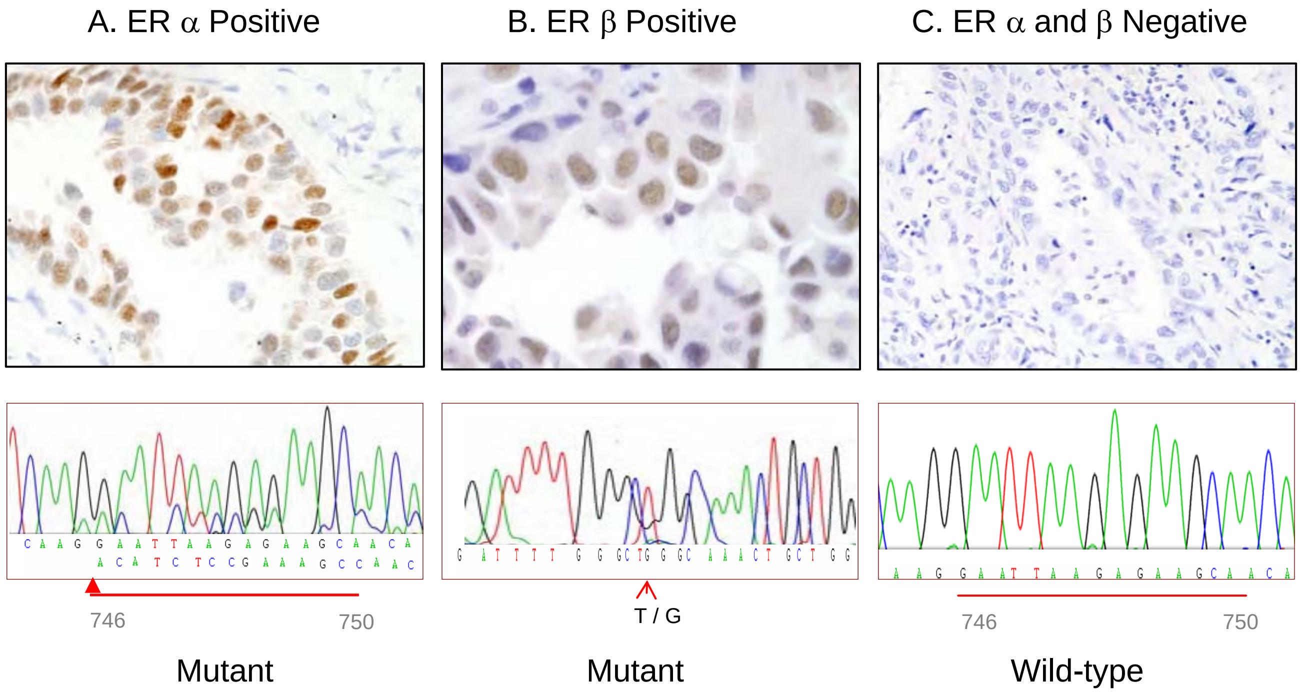


Figure 3



A. ER α Positive
B. ER β Positive
C. ER α and β Negative
C A A G G A A T T A A G A G A A G C A A C A
A C A T C T C C G A A A G C C A A C
746
750
Mutant
G A T T T T G G G C T G G G C A A A C T G C T G G
T / G
Mutant
A A G G A A T T A A G A G A A G C A A C A
746
750
Wild-type

TTF-1 Gene Amplification and Protein Expression Pattern Identify Adenocarcinoma of Lung with Worse Prognosis.

Ximing Tang[1], Menghong Sun[1], Carmen Behrens[1], Ludmina Prudkin[1], Natalie Ozburn[1], Adi F. Gazdar[3], Cesar Moran[1], Marileila Varella-Garcia[2], and Ignacio I. Wistuba[1]
[1]UT MD Anderson Cancer Center, Houston, TX 77030; [2]University of Colorado, Aurora, CO 80011; [3]UT Southwestern Medical Center, Dallas, TX 75235

Thyroid transcription factor -1 (TTF-1), a lineage-specific transcription factor frequently overexpressed in lung adenocarcinoma, has been recently reported to show gene amplification in a subset of these tumors. To better characterize *TTF-1* copy number in non-small cell lung carcinomas (NSCLC) and correlate it with protein expression, we studied both gene copy number and protein expression using fluorescent *in situ* hybridization (FISH) and immunohistochesmistry stain (IHC) assays in a large series (N=324) of surgical resected NSCLCs placed in tissue microarrays, including 205 adenocarcinomas and 119 squamous cell carcinomas. We correlated our findings with patients' clinico-pathologic characteristics, and in a subset of adenocarcinomas with *EGFR* (exons 19-21) and *KRAS* (exons 1 and 2) mutation status. *TTF-1* amplification (FISH$^+$, clustered gene signals) was detected in 19% (51 out of 269) of tumors, without differences by histology (18% of squamous cell carcinoma and 19% of adenocarcinoma). TTF-1 protein high level expression (IHC$^+$, semiquantitative score ≥200, range 0-300) was detected exclusively in adenocarcinomas (48% of cases), and in this tumor type correlated with gene amplification (P=0.005). No correlation between TTF1 abnormalities and patients' age, gender, smoking status and TNM stages was detected. In adenocarcinomas, IHC$^+$, but not FISH$^+$, correlated with *EGFR* and *KRAS* mutations: IHC$^+$ was more frequently found in *EGFR* (16/21, 76% vs. 59/172, 34%, P<0.001) and *KRAS* (8/11, 72% vs. 26/75, 34%, P=0.016) mutant compared with wild-types tumors. Survival analysis showed that for adenocarcinoma *TTF-1* FISH$^+$ correlated with worse recurrence-free survival (P=0.001), while IHC$^+$ correlated with better recurrence-free survival (P=0.036). Our findings indicate that *TTF-1* amplification occurs in a subset of NSCLCs, including both major tumor histologies: adenocarcinoma and squamous cell carcinoma. The association of TTF1 expression with *EGFR* and *KRAS* mutation in lung

adenocarcinomas may correlate with the peripheral airway origin of these tumors. Both TTF1 gene amplification and protein expression correlate with NSCLC patients' prognosis. (Supported by Grant DoD-W81XWH-04-1-0142 and W81XWH-05-2-0027).

TITF-1 and *EGFR* gene copy variations are associated with prognosis for the patients with non-small cell lung cancer

Ximing Tan, Diane Liu, Carmen Behrens, Dandan He, Menghong Sun, David Rice, J. Jack Lee, Waun K. Hong, and Ignacio I. Wistuba
UT MD Anderson Cancer Center, Houston, TX 77030

Thyroid transcription factor -1 (TITF-1, a lineage-specific transcription factor), and the epidermal growth factor receptor (EGFR, a tyrosine kinase membrane receptor) have shown frequent gene amplification in non-small cell lung carcinomas (NSCLC). We investigated the clinico-pathologic characteristics of NSCLCs having *TITF-1* andor *EGFR* gene copy number abnormalities by examining gene copy number status using quantitative polymerase chain reaction (qPCR) and DNA extracted from microdissected formalin-fixed and paraffin-embedded tissue from 53 adenocarcinomas and 29 squamous cell carcinomas. ß-actin gene was used as reference. In tumors, gene copy ratio referred to *ß-actin* ranged from 0.22 to 74.93 (median=1.52) for *TITF-1*, and 0.05 to 6.28 (median=1.51) for *EGFR*. Ratios 1 to 2 were defined as normal gene copy number (NGC). Ratios <1 and >2 were defined as low gene copy (LGC) and high gene copy (HGC) number, respectively. Both, LGC and HGC categories were defined as abnormal gene copy. Similar frequencies of *TITF-1* and *EGFR* copy number categories were detected comparing adenocarcinoma (*TITF-1*: LGC 15, 28%; NGC 20, 38%; HGC 18, 34%; *EGFR*: LGC 9, 17%; NGC 27, 51%; HGC, 17, 32%) with squamous cell carcinoma (*TITF-1*: LGC 8, 28%; NGC 14, 48%; HGC 7, 24%; *EGFR* : LGC 5, 17%; NGC 18, 62%; HGC 6, 21%). We found a statistically significant correlation between *TITF-1* and *EGFR* copy numbers (Spearman correlation coefficient=0.36, *P*=0.0008). In both tumor histologies, neither *TITF-1* nor *EGFR* gene copy increase (ratio >2) correlated with disease prognosis**.** However, in adenocarcinomas, Kaplan-Meier and log rank tests revealed that the median time to death was longer in patients with normal copy number compared with those with abnormal copies for *TITF-1* (median 4.76 years, 95% CI 2.95~NA, *P* = 0.04) and *EGFR* (4.76 years, 95% CI 3.13~ NA, *P*=0.04). Moreover, adenocarcinoma patients with combined *TITF-1* and *EGFR* abnormal copy showed worse overall survival (3.56 years, 95% CI 3.13~ NA) compared with patients with normal copy

status (median not reached, *P*=0.003). In these patients, multicovariate Cox modeling indicated that combined copy abnormality of both genes is an independent factor for worse overall survival (HR 4.566, *P*=0.0057). Our findings suggest that loss and gain of *TITF1* and *EGFR* are frequent abnormalities in both adenocarcinomas and squamous cell carcinomas of the lung, and in adenocarcinoma patients correlate with disease outcome. (Supported by Grant DoD-W81XWH-04-1-0142 and W81XWH-05-2-0027).

(*TITF-1* and *EGFR* gene copies in lung cancer)